

# Foreign Vocabulary in Classical Arabic and al-Jawālīqī's *al-Mu'arrab*

Alex Boysen



## **ARA4590: Master's Thesis in Arabic Studies**

Thesis submitted in partial fulfilment (60 credits) of a two-year (120 credits) degree in the field of Arabic language

Programme for Asian and African Studies  
Department of Culture Studies and Oriental Languages

**University of Oslo**

May 2009



*That proposition which is especially beloved...  
is that all should converse in the Arabic  
language.*

*This, inasmuch as it is the most comprehensive  
(‘absaṭ) of all languages. ...*

*The Persian language is extremely sweet. ...  
Persian, however, does not, and will never have,  
the magnitude of Arabic.*

*Indeed, relative to it, all languages have been,  
and will remain, circumscribed.*

*- Bahá’u’lláh (1817-1892)*

# Foreign vocabulary in Classical Arabic and al-Jawālīqī's *al-Mu'arrab*

## Table of Contents

Abbreviations	7
<b>Introduction</b>	<b>8</b>
<b>1. The process of codifying the Qur'ān text</b>	<b>16</b>
1.1 Initial steps – the appearance of a class of learned men	16
1.2 The pioneering work of Muqātil – categorizing of the holy text	17
1.3 Methods and conceptual tools for analysing the holy text – the development of a specialized technical-descriptive vocabulary	17
1.4 The practice of <i>tafsīr</i> versus linguistic analysis: scientific endeavors begin to take on a life of their own	18
<b>2. The process of codifying the collected corpus of the Arabic language</b>	<b>20</b>
2.1 The beginnings of linguistic science – means of transmission	20
2.2 The first linguists – their concerns, activities, publications and place within the Islamic scientific enterprise	22
2.2.1 The construction of grammar: creation of basic categories	22
2.2.1.1 Nouns as elemental - distinguishing indigenous and non-indigenous modes of expression	23
2.2.1.2 Verb as derivative of noun, indicative of process and time	24
2.2.1.2.1 Importance of derivation in Arabic grammar	24
2.2.2 Abū l-Aswad's foundational role	26
2.2.3 Twofold specialization merging in the subject of derivation	27
2.2.4 Concerns of the early linguists	28
2.2.5 The earliest word lists	30
2.2.6 First major systematizations of the language	31
2.2.7 Relationship of the linguistic disciplines to Qur'ān exegesis	32
2.3 Al-Khalīl ibn Aḥmad and his major work, <i>Kitāb al-'Ayn</i>	33
2.3.1 <i>Kitāb al-'Ayn</i>	34
2.3.1.1 Historical assessment, scope and formative influences	34
2.3.1.2 Generative conception	35
2.3.1.3 Organization of entries	36
2.3.1.4 Constitutive theory and basic intention	38
2.4 Subsequent developments in lexicography	39
2.5 Sībawayhi	40
2.5.1 The <i>Kitāb</i>	40
2.5.1.1 Contents and influences	41

## Table of Contents

2.5.1.2 Sources for his work	42
2.5.1.2.1 <i>Samā</i> ‘attested data’	42
2.5.1.2.2 Scope of his research	43
2.5.1.3 Analytical concepts and methodological approaches	44
2.5.1.3.1 Methodological considerations related to <i>ši’r</i> and <i>kalām</i>	44
2.5.1.3.2 The concept of analogy ( <i>qiyās</i> )	45
2.5.1.4 Characterizing the morphology and phonology of words in Arabic	46
2.5.1.4.1 Morphology	46
2.5.1.4.1.1 The root system of Arabic words	46
2.5.1.4.1.2 Morphological patterns of words	49
2.5.1.4.1.3 Syllable structure and its consequences for morphology	49
2.5.1.4.1.4 Other morphological processes that can potentially affect loan words	50
2.5.1.4.2 Phonology	51
2.5.1.4.2.1 Changes occurring in the phonology of loan words	51
2.5.1.4.2.2 Basic assumptions regarding phonology	52
2.5.1.4.2.2.1 Disallowed or disapproved sound combinations in Arabic	53
2.5.1.4.2.2.2 Tracing sound changes in loan words	55
2.5.1.4.2.3 Method of describing Persian phonemes	57
2.5.2 Speech as an expression of behaviour	57
<b>3. Factors affecting the process of canonization of the Arabic language</b>	<b>58</b>
3.1 The role of the Qur’ān and Islamic institutions in the canonization of the language	58
3.2 Social factors influencing the canonization of the language	61
<b>4. The role of identifying foreign vocabulary in the process of canonizing the language</b>	<b>64</b>
4.1 Perspectives on and criteria for distinguishing foreign vocabulary in Arabic	64
4.1.1 Attitudes of the learned to the question of the existence of foreign vocabulary in the Qur’ān and in the Arabic language in general	64
4.1.1.1 Muslim grammarians’ reckoning of foreign vocabulary in Arabic	66
4.1.1.2 Theological hegemony and the eventual penetration of secular ideas into the Islamic domain	68
4.2 Al-Jawālīqī, author of <i>al-Mu‘arrab</i>	69
4.2.2 <i>Al-Mu‘arrab</i>	69
4.2.2.1 Contents	70
4.2.2.1.1 Form and approach	70
4.2.2.1.2 Subject matter	70
4.2.2.1.3 Standpoints on the issue of the occurrence of loan words in the Qur’ān	71
4.2.2.1.4 Reflections over the place and function of loan words in the language	72
4.2.2.1.5 Ordering of the entries	73
4.2.2.1.6 Transcription of supposed loan words	74
4.2.2.2 Al-Jawālīqī’s treatment of <i>samā</i>	74
4.2.2.3 Analytical concepts and methodological approaches	75
4.2.2.4 Themes	75
4.2.2.4.1 Plants, fruits, animals	75
4.2.2.4.2 Christians	76
4.2.2.4.3 Certain “un-Islamic” subjects	77

## Table of Contents

4.2.2.4.4 Prophets and names figuring in the Qur'ān	77
4.2.2.4.5 Application of miscellaneous grammatical rules	79
4.2.2.4.6 Personages, peoples, languages, common names	80
<b>Conclusion</b>	<b>83</b>
<b>Appendix 1</b>	<b>85</b>
<b>Appendix 2</b>	<b>88</b>
<b>Bibliography</b>	<b>90</b>
<b>Abstract</b>	<b>96</b>

## **Abbreviations**

EALL	Encyclopedia of Arabic Language and Linguistics
EI	Encyclopaedia of Islam
ELL	Encyclopedia of Language and Linguistics
EQ	Encyclopedia of the Qur'ān

# Introduction

---

The present thesis was written over two semesters as part of my degree in Arabic at the University of Oslo. My interest in the subject of foreign words in Classical Arabic was first sparked by a recommendation by my supervisor, Professor Lutz Edzard of the University of Oslo, that I investigate scholarly work done on the major sourcebook for such vocabulary in the history of Arabic linguistics, al-Jawālīqī's *al-Mu'arrab min al-kalām al-'a'jamī 'alā ḥurūf al-mu'jam*, literally '[What has been] Arabicized from foreign speech, arranged according to the letters of the alphabet'. A shortened version might read, simply 'Arabicized foreign speech'.

A note on the terminology used throughout to designate foreign vocabulary is in order: the expressions 'foreign vocabulary,' 'loan vocabulary' and 'loan words' are used interchangeably; a fourth alternative is the simple noun 'borrowing'.

The study presented here traces significant historical developments in Arabic linguistics leading to the canonization of the language, and the place of the question of distinguishing foreign vocabulary in the context of these developments, with reference to al-Jawālīqī's work. It goes without saying that the latter, specific problem is inseparable from the wider historical context of the endeavors of Muslim linguists to distinguish the normative features of Arabic. The issue of foreign vocabulary is thus, strictly speaking, not a subject per se; rather, it must be viewed as an outcome of the larger process of clarifying the distinguishing features of the Arabic system itself. In answer to this, the present thesis has been structured around the main features of the said historical developments that have proved particularly relevant for differentiating indigenous from foreign words, relating these as far as possible to the contents of *al-Mu'arrab*. It is left to the reader to judge the relevancy of the operation thus performed. The latter work, which is the unmatched compilation of instances of supposed foreign words in Arabic and reports thereon, has been of the greatest assistance as a sourcebook by which to understand and gauge the issue at hand. In light of its importance to the question of loan words in the language, I have made it the object of a case study as well.

Regarding the Arabic grammatical tradition, Carter (1999) has synopsized its major developments in a four-stage chronology which, although related specifically to the process of



## Introduction

defining the general rule (*qiyās*) in relation to the individual exception (*šādd*) in grammar in the so-called Baṣran-Kūfan debate, provides an elegant introduction to the tradition as such, and an added parameter by which developments related to the question of loan vocabulary may be analyzed. The period surveyed herein covers the first two stages corresponding roughly to sections 2-3.2 and 3.3-5.2 respectively. The issue of loan words will be seen to have been given systematic treatment only in the course of the second stage.

In the first stage, which can be dated to much the same time as the early collections of pre-Islamic poetry and attempts at Qur'ānic exegesis, i.e. the first half of the eighth century, the emphasis is on gathering linguistic material, with almost no processing or analysis. The motivation is to preserve the records of the vanishing past, the operating principle is honesty, diligence and accuracy, and the output is raw data.

A second stage, represented by Sībawayhi in the second half of the eighth century, recognizes the systematic nature of language. Building on the data already collected, a huge exercise of pure induction is carried out in a deliberate effort to survey the entire known language, and all the material is then distributed into linguistic categories. The operating principle is self-conscious analogy both by speaker and observer, and the output is an exhaustive and systematic description of Arabic... (Carter 1999: 66)

Carter's designation of "analogy" as the "operating principle" in the latter stage will be seen to carry particular relevance for the subject of this thesis.

As to the work of the lexicographer, Haywood (1965: 102) presents the argument that dictionary writing "is only for hacks," presumably in the sense of being a discipline essentially attendant to the activities and outcomes of the more "innovative" branches of linguistic science such as grammar and literature. Even if this were true to some extent, it can in no way diminish the importance of this branch of study to the system of linguistics in general as demonstrated by the sustained and serious attention afforded it by some of the brightest scholars in Islam, notably al-Khalīl ibn Aḥmad (d. c. 175/791) (perhaps the greatest of them all, said to have written the first Arabic dictionary) and men such as Ibn Fāris (d. 390/1000), al-Jawharī (d. c. 398/1007) (credited with inventing the rhyme arrangement in Arabic lexicography), Ibn Manẓūr (d. 711/1311) and al-Fīrūzābādī (d. 817/1414). The idea (of the work of a hack) brings into perspective as well the cumulative nature of knowledge—the scientific principle—the advancement of which relies as much on the work done by one's

predecessors in the field as on what Partridge (1963: 38) labels, in our case, the “hard work” of the “conscientious lexicographer,” which he contrasts with the supposedly more leisurely work of “merely adapting someone else's dictionary.”

Such considerations bring us back to one of the challenges met with in the course of preparing the present study due to the largely undifferentiated nature of the theme being discussed: al-Jawālīqī has certainly done the “hard work” that producing a dictionary of foreign vocabulary in Arabic involves viz. the collection and organizing of pieces of information relevant to the subject. There has, moreover, been little to expand upon or to elaborate in this regard, as the fact of *al-Mu‘arrab*’s not having been superseded by any comparable work to this day clearly testifies. As the present study will show, the latter work has presented the undersigned with limited scope for operationalizing (in the sense of, “defining a concept or variable so that it can be measured or expressed quantitatively”<sup>1</sup>), probably due to the fact of it representing perhaps the lowest level in the empirical hierarchy; the most easily grasped, thoroughly digested and readily available elements in linguistic science: words in and of themselves. Kopf (1961: 197) has stated poignantly with reference to the subject of tracing the origin of foreign words: “Aside from phonetic considerations, no proper scientific principle is discernible in this field of lexicographical research.”

It may also be noted in this connection that foreign vocabulary does not express a uniform category but rather a motley group of seemingly haphazard selections from the language ranging widely in theme and externals. In researching the present topic, these very characteristics have actuated an exploration of probably a wider range of themes in linguistics than would normally have been required for the writing of a Master’s thesis in Arabic, thus providing an eclectic research base for anticipated continued investigations in the present and related fields.

As an aid to conceptualizing the subject under examination it will be helpful to dwell somewhat on the theme of foreignness itself. In the Qur’ān, kindness to the foreigner is explicitly enjoined, a principle that in practice would lessen the inevitable barrier to association and intercourse that differences of language, culture and outlook represent, besides sheer distance:

---

<sup>1</sup> "operationalize." Webster's New Millennium™ Dictionary of English, Preview Edition (v 0.9.7). Dictionary.com, LLC. 17 May. 2009. <Dictionary.com  
<http://dictionary1.classic.reference.com/browse/operationalize>>.

## Introduction

*wa-‘budū llāha wa-lā tušrikū bi-hi šay’an wa-bi-l-wālidīna ‘ihsānan wa-biḍī l-qurbā  
wa-l-yatāmā wa-l-masākīni wa-l-jārī ḍī l-qurbā wa-l-jārī l-junubi wa-l-sāḥibi bi-l-  
janbi wa-bni l-sabīli wa-mā malakat ‘aymānukum ‘innā llāha lā yuḥibbu man kāna  
muḥtālan fa-ḥūran \* ‘alladīna yabḥalūna wa-ya’murūna l-nāsa bi-l-buḥli wa-  
yaktumūna mā ‘ātāhumu llāhu min faḍlihi wa-‘tadnā li-l-kāfirīna ‘adāban muḥīnan  
(Q 4:36-37)*

Serve God, and associate naught with Him. Be kind to parents, and the near kinsman, and to orphans, and to the needy, and to the neighbour who is of kin, and to the neighbour who is a stranger, and to the companion at your side, and to the traveller, and to that your right hands own. Surely God loves not the proud and boastful \* such as are niggardly, and bid other men to be niggardly, and themselves conceal the bounty that God has given them. We have prepared for the unbelievers a humbling chastisement... (Arberry)

*‘innā llāha lā yaḍlimu miṭqāla ḍarratin wa-‘in taku ḥasanatan yuḍā‘ifhā wa-yu’ti min  
ladunhu ‘ajran ‘aẓīman (Q 4:40)*

Surely God shall not wrong so much as the weight of an ant; and if it be a good deed He will double it, and give from Himself a mighty wage. (Arberry)

A plausible anecdote, though referring to Bedouins of the early twentieth century but with obvious roots to a more distant past, states: “Among the Arabs there is no better report of a man’s life than to be called in his country *karīm*, a liberal soul; so nothing more hateful than the lean niggard’s name, *baḥīl*” (Doughty 2000: 430). It is not clear, however, whether this outlook would apply as much to foreigners as to those of one’s own kind. A more unequivocal statement of the Bedouins in this regard has it that, the guests are “guests of God” (Westermarck 1924: 580).

The attitude of the pre-Islamic Arabs to foreigners can only be guessed at from the accounts related about them and poetry surviving from the period. The threefold themes of such poetry were *madīḥ* ‘eulogy,’ involving praise of the living; its antithesis *hijā* ‘invective/lampoon;’ and *riṭā* ‘elegy,’ praise of the dead. All three served to celebrate the values of the community, either directly through praise or indirectly by highlighting the virtues of one’s own community in relation to the vices of others in the form of lampoon (Allen 1998: 138-9). History has shown that the teachings of the Qur’ān represented a clear

## Introduction

break with many of the communal values of the pre-Islamic Arabs, one example being the degree of tolerance and justice extended to foreigners as illustrated in the above Qur'ānic passages. One may surmise on this basis that the door to the adoption of foreign vocabulary may well have opened wider with the coming of Islam, but it would be difficult to draw any reasonable conclusions about the changes in receptivity of the Arabic language to such vocabulary from the little we know about the age of *jāhilīya*. The Arabic term '*ajam*' refers to the "people qualified by '*ujma*', a confused and obscure way of speaking, as regards pronunciation and language," its antithesis being *faṣāḥa* 'purity of the language'. The '*ajam*' are also the non-Arabs. Pre-Islamic poetry includes the contrasting of '*ajam*'—referring to their immediate neighbors the Persians—with '*arab*'. Rodinson outlines the dual implications of the term in this context:

The affective value attributed to the word depended on the point of view of the user; although it preserved for the most part the original contemptuous force inspired by the haughty presumptuousness of Arab superiority, it sometimes, and even at an early date, implied the desirability and allurements of the exotic, and the acknowledgement of a more civilized and refined culture. (Rodinson EI Vol. I: 205)

It is not unlikely that attitudes such as these, implicit in '*ajam*', should apply in part to the issue of foreign vocabulary as well, although what this can have entailed in practice for the dynamics of the adoption of loan words in Arabic will remain a matter of conjecture.

The technical term in linguistics used to designate the phenomenon of foreign vocabulary is 'borrowing,' defined by Crystal (1993: 46) as "the introduction of a word (or some other linguistic feature) from one language or dialect into another." Heath defines it as "a form that has spread from one linguistic variety (the 'source') into another variety (the 'target' or 'replica')." He regards 'loan word' (or 'loanword') as nearly synonymous to borrowing, the difference being that the latter "is often really a stem (smaller than a word), and may be a phrase (larger than a word)," and adds that "[b]orrowing is also the term for the act of incorporation itself" (Heath ELL 383). Another definition describes borrowing as, "Adoption of a linguistic expression from one language into another language," and adds with reference to factors that influence its occurrence, "usually when no term exists for the new object, concept, or state of affairs" (Bussmann 1996: 55). The latter judgment expresses a significant particularity of foreign vocabulary that has received little attention in the case of

Arabic, likely owing to the complications involved with tracing etymologies to a distant past. Bussmann goes on to suggest possible motives for this phenomenon:

Among the causes of such cross-linguistic influence may be various political, cultural, social or economic developments (importation of new products, prestige, local flavor, internationalization of specialized languages and jargons, among others). (Bussmann 55-56)

This general evaluation is certainly pertinent to the case of Arabic as well, as the survey of certain themes found in *al-Mu‘arrab* in section 4.2.2.4 will serve to demonstrate.

As to the place of the subject of borrowing within the larger scheme of linguistics—reflecting, besides, on the issues being discussed in the present thesis—Heath writes:

The study of borrowings is of interest to general linguistics because the borrowing language may have several possible ways of incorporating the foreign form into its own phonological, morphological, and semantic systems, and the options implemented may reveal something about deep-seated developmental tendencies of the language that are not otherwise clearly evident. (Heath 383)

Although such “deep-seated developmental tendencies” in Arabic are not a subject in this study, I believe the themes covered here can certainly shed light on these tendencies by way of the more “superficial” themes being treated. These include the ideological and historical circumstances and personages that have shaped Classical Arabic as we know it, as well as practical aspects affecting the adaptation of loan words such as morphological and phonological changes that occur in borrowing besides specifics of actual cases of it. Referring back to Heath, the present study relates probably more closely to his characterization of borrowing patterns as reflecting “the social and historical context in which the language contact takes place” (Heath 383).

Besides general research on the history of Arabic grammar, the current study has involved translation of relevant passages from al-Jawālīqī’s *al-Mu‘arrab* for citation herein. I have used the critical edition of *al-Mu‘arrab* prepared by Aḥmad Muḥammad Shākir as the basis for these translations as it is reliably based on four manuscript copies of the work (al-Jawālīqī 1969: 21). Where discrepancies existed between the different manuscripts, the editor put the parts missing in some of the versions in brackets. In the translations that follow I have

## Introduction

opted to include all the material that was put in brackets by the editor thus making the brackets superfluous to the reader, for which reason I have removed them in the translation. I have sought to make the translation as literal as possible; certain words and names have been inserted in brackets in order to facilitate an understanding of the context wherever it was thought necessary, and a transcription of the original Arabic word has occasionally been added in parentheses after the translated word for this same reason.

Arabic-English dictionaries I have used in making the translations are, in order of use, those of Wehr (Cowan 1994), Hava (1899) and (very occasionally) Lane (1968). Hava's dictionary has been found especially helpful since it is based on Classical Arabic, covering the period of al-Jawālīqī's work. The specialized character of parts of al-Jawālīqī's work, besides its overall length, at a point impelled the need to rationalize the amount to be translated from it, especially the poetry citations, which frequently employ rare and dated terminology that is not listed in Hava's or Wehr's dictionaries. Accordingly, certain vocabulary items have thus been merely transcribed rather than translated. Correct and faithful translation from Arabic into English is generally no easy task, and there are bound to be errors in the translations given here. For these I take the sole responsibility, and sincerely apologize for them. Three consecutive stops [...] indicate where parts have been omitted.

I have not found a project of this specific type to have been undertaken before, and so it is my hope that the present study can help clarify the main historical developments that gave rise to the concept of Arabic as we know it in relation to the subject of foreign vocabulary, and give a glimpse of the various issues that arose in relation to the problem of identifying loan words in the language.

I must thank my supervisor Professor Edzard for his excellent advice and guidance in the course of preparing this thesis, without which it would not have seen the light of day! Thanks go also to my learned friend Amund Bjørnsnøs for his kind and helpful suggestions on scholarly matters and in relation to the subject.

## Notes on transcription

Two different standards have been used for transcribing the sounds /th/, /kh/, /sh/ and /gh/ respectively: [t̪], [ħ], [ʃ] and [g̤] has been used for transcribing Arabic text (italicised), while [th], [kh], [sh] and [gh] is used in other cases, as well as in names (except those dealt with as

## Introduction

cases of supposed foreign vocabulary in *al-Mu‘arrab*). The transcription of the Arabic verb يسهل ‘to make easy’ becomes, for instance, *yashula*.

## The process of codifying the Qur'ān text

*Practical and empirical developments in the process of systematizing study of the holy book.*

---

As we will see in the following, the coming of Islam and the appearance of the Qur'ān marked the inception of prodigious endeavors in learning of every kind, notably in the area of linguistics under consideration here. The growth of Islamic science is a fascinating and instructive subject in itself, offering important lessons on the processes of civilization.

### 1.1 Initial steps – the appearance of a class of learned men

The initial stage in the process of codifying and analyzing the Qur'ān consisted in its collection after the Prophet's passing, a work that required a high level of linguistic and technical expertise for the various tasks involved, such as the reform of orthography, the evaluation of variant readings, the clarification of unusual forms, and the prioritizing of dialectal variants (Versteegh 1997: 8-9). These efforts opened the way for the establishment of a group of learned men who could elucidate obscure, mysterious and complex passages and ordinances of the holy text and apply its guidance to situations of everyday life in the community. In exercising Qur'ān exegesis (*tafsir*), arguably the first scientific discipline in Islam, the earliest commentators largely emphasized the *meaning* of the text and did not discuss divergences in grammar that existed between different readings (*qira'āt*). The focus of study was still on the problems surrounding practical aspects of daily life, to which they sought answers from God's word. Muqātil is a representative of the first generation of this newborn exegetical class which, as a result of the rapid expansion of the Islamic empire, came to play an increasingly significant role in the order of Islamic society.



## 1.2 The pioneering work of Muqātil – categorizing of the holy text

Muqātil ibn Sulaymān's (d. 767) Qur'ān commentary summarizes the subject matter of the holy book as follows, describing its contents in a manner that will be seen to relate to later developments in the technical-linguistic analysis of the text:

Muqātil said: The Qur'ān contains references to particular and to general things, particular references to Muslims, and particular references to polytheists, general references to all people. It contains ambiguous and univocal passages, explained and unexplained passages; it contains deletions and explicit utterances; it contains connective items; abrogating and abrogated verses; it contains changes in the chronological order; it contains similar utterances with many different aspects; it contains passages that are continued in a different *sūra*; it contains accounts of what is in the hearts of the believers, and accounts of what is in the hearts of the unbelievers, polemics against the Arabian polytheists, and it contains explanations, and for each explanation there is an explanation (Versteegh 1997: 11-12).

This succinct categorical synopsis presaged a wealth of scholarly endeavors that scrutinized every aspect of the language, focused to begin with on the holy text in particular and later expanding to embrace the Arabic language in general.

## 1.3 Methods and conceptual tools for analysing the holy text – the development of a specialized technical-descriptive vocabulary

The dialectic milieu that arose in response to the initial efforts to understand and apply the provisions of the holy book later diversified into branches that treated specifically its technical and linguistic aspects, such as the formal characteristics of words and sentences, morphology, phonology, grammar, etymology, dialectal origins, etc. Even though the structure of the language of revelation was not an explicit point of interest to the early commentators, Muqātil's work does include digressions such as occasional remarks on the etymology and dialect of words, as well as elaboration of certain terminology such as

“deletions” and “connections” as found in the above passage. For instance, these latter are closely related to his use of the technical terms *iḍmār* ‘hiding’, and *ṣilāt fī l-kalām* ‘connections in speech’ respectively, which is a development into specialized meanings that can be seen to herald the beginnings of what became an elaborate and systematic structural and formal analysis of the Qur'ān, albeit still born of an essentially semantic approach.

The early Qur'ān commentaries drew on a very limited technical vocabulary, and the same term would be used in its vague meaning to describe otherwise incongruent phenomena. An example is the device of *taqdīm* ‘preposing’, which Muqātil uses to designate three things: what we today would call *hysteron proteron* (a change in the logical order of events), prolepsis (when the result of an action is presented as coexisting with it), and syntactic hyperbaton (a change in word order). Thus his use of *taqdīm* does not distinguish between linguistic and semantic textual analysis, and only later acquired its present technical meaning in syntax of fronting (Versteegh 1997: 11-12).

#### **1.4 The practice of *tafsīr* versus linguistic analysis: scientific endeavors begin to take on a life of their own**

The process of canonization of the Qur'ān and *aḥadīth* that came to facilitate further systematic study of the meaning and import of the message of Islam had been completed in the course of the Umayyad period (660-750 C.E.) (Seidensticker EALL Vol. III: 30). As we have seen, the specific practices of recording, editing, preserving, transmitting, monitoring, interpreting, and teaching the text of the Qur'ān opened the way for the more advanced discipline of *tafsīr* (exegesis), which focused primarily on its semantic features (Carter 1998: 31). In the early stages of the discipline there were no branches of specialization focusing each on a different aspect of the Qur'ān text and message, and *tafsīr* works thus presented an amalgam of Islamic scholarship dealing with such disparate fields as historical narrative, abrogation, pre-Islamic folklore, lexicography, legal application, theology, semantics and grammar (Versteegh 1993: 195). These aspects were emphasized in varying degrees by the *mufasssīrūn* ‘exegetes’ according to their respective goals and concerns.

Linguistic and grammatical studies complemented the *tafsīr* discipline and widened its base by including under its purview the pre-Islamic poetry and *šawāhid* from Arabic dialects. These corpora made up the formal linguistic ground from which the “Arabic” holy book and the *sunna* accounts had emerged, and would serve to verify the true lexicographical meaning

of words occurring in the Qur'ān and the *aḥadīṭ* (Seidensticker EALL Vol. III: 32). In general, the earliest period in the development of the Arabic grammatical tradition—represented principally by Sībawayhi—was characterized by an emulation of the norms prevailing in the initial phases of Arab history, relating as much to law, ethics and aesthetics as to the language itself (Carter 1973: 146). In fact, such a preoccupation can be said to have characterized the entire course of the discipline: The sum of the exercises involved in preserving the sanctified heritage of the holy text for posterity, and, by extension, of the Arabic language system itself as seen in the Qur'ān's recitation, transmission and teaching, has been understood by Carter (1998: 31) in terms of the religious practice of perpetually reproducing God's singular Act of revelation. Such a retrograde orientation contrasted with the immediate legal- and practical oriented discipline of *tafsīr*, whose task was to relate the formulae of the Qur'ān to ever-changing trends of thought and to novel situations in the life of society. The former was directed exclusively to historical data to which was attributed eternal validity; the latter was constantly confronted with new circumstances and problems to which it would be obliged to provide relevant pronouncements and solutions (though not without recourse to befitting precedents!).

## **The process of codifying the collected corpus of the Arabic language**

*Key movers and developments effecting the systematization of the language.*

---

### **2.1 The beginnings of linguistic science – means of transmission**

The insight and knowledge held by the scholars in Islam was at first scattered and unsystematized due to the oral nature of its transmission, the former fact preventing a high degree of specialization. There was thus in early Islam a considerable interplay between the various scholarly disciplines and almost all the grammarians were occupied with more than one field apart from grammar (Versteegh 1993: 192).

In the context of oral transmission knowledge was systematized in accordance with the particular exigencies of this mode of communication, characterized by Fox (2006) as follows:

Orality depends upon a real or simulated ‘live’ and copresent interaction. Sound emerges and dies away in the instant of communication, gesture in the blink of an eye. Its inscription freezes some essence of meaning, making meaning available in new and unforeseen contexts of use and interpretation. ...meaning of discourse depends on a calibration of shared social experience in the fleeting moment of instantiation, and...some kinds of meaning are greatly enhanced and amplified under these conditions, just as others are diminished or obscured. (Fox ELL 82)

The following survey by Daniels (2001) of the history and role of the spoken medium of communication demonstrates the primacy of the oral medium as a factor in the development of the Arabic language:

Writing is indispensable for civilization – but entirely irrelevant for language. Most of the thousands of human languages were never written until recent years, and their speakers were none the worse for it. Their cultures were full and rich, lacking only accountancy and science. Everything else that is written need not be: poetry, narrative, law, and their apotheosis, scripture, are all part of every oral culture. Only in a city is the community so large that letters must be sent to communicate personal messages... Cities are where production does not link directly with consumption...

But cities characterize only a handful of human societies, and the vast majority of human languages never had written forms of their own. The discovery that languages other than the classical ones were every bit as rich as Greek, Sanskrit and Chinese...led linguists to concentrate on unwritten languages and then to devalue the study of written records in favor of fieldwork. (Daniels 2001: 75)

Reflections such as these reveal the relative lack of a conceptual apparatus to distinguish the ‘indigenous’ from the ‘foreign’ when it came to vocabulary origins in Arabic before the coming of Islam, at which the appearance of civilization and a written system for the language generated the intellectual instrumentalities required for such differentiation. In relation to foreign vocabulary, it is uncertain how oral and written cultures respectively would influence receptivity to and rates of adoption and retention of foreign vocabulary in Arabic, or whether there would be any difference at all in this regard, although Fox’s discussion of orality in terms of its potential to transmit different types of *meaning* (with *form* being a possible counterpart to the latter as seen in the traditional dichotomy between *ma’nā* and *lafẓ* ‘form’ in Arabic linguistics) could conceivably be used as a prospective entry point for the discussion. Foreign vocabulary is usually differentiated by its atypical morphological features—its *lafẓ*—while *ma’nā* is generally less reliable as an indicator of foreignness (obvious exceptions being phenomena specific to a foreign people, e.g. *majūs* ‘magician’ which is specific to Persians, from Middle Persian *magapat* ‘the chief of the Magi, i.e. the main priest of the Zoroastrian clergy,’ or Old Persian *magul* [Asbaghi EALL Vol. III: 581]). With reference to loan words entering into Arabic from Persian, a potential indicator for the rate of adoption could be the fact (as presented by Asbaghi [580]) that the process of borrowing increased with the spread of Islam as a result of the uniting of peoples of different backgrounds.

Apropos Persian, the designation is used by Muslim grammarians to refer to the later Persian language they were acquainted with, and not Pahlavi (termed Middle Persian), the Persian language spoken in pre-Islamic times and the official language of the Sasanid Empire

(226-640 C.E.). This discrepancy highlights the many complications of determining the exact etymological path of supposed loan words from Persian in Arabic (and probably from other languages as well) (Jeffery 1938: 15-16).

## **2.2 The first linguists – their concerns, activities, publications and place within the Islamic scientific enterprise**

### **2.2.1 The construction of grammar: creation of basic categories**

The following anecdote about Abū l-Aswad al-Duʿalī (d. ca. 69/688), who is generally acknowledged to be the first Arabic linguist, highlights the significance of the simple tripartite division of the Arabic language into parts of speech, which would have such far-reaching effects on the development of the whole course of the Arabic grammatical tradition.

According to Abū ʿUbayda (d. 209/824-5), Abū l-Aswad had learned grammar from the Caliph ʿAlī ibn Abī Ṭālib, Muḥammad’s cousin and son-in-law. The latter is said to have handed him a manuscript in which was written: “Language is noun and verb and particle. The noun is what informs about a named object; the verb is that with which the information is given; and the particle is what comes for a meaning.” He was then told: “Follow this direction (*naḥw*) and add to it what you find!” (Versteegh 1993: 4.) Another very similar version has the manuscript text as stating simply: “The noun (*ism*) is what describes what has a name; the verb (*fiʿl*) is what describes the movement of that which has a name; the particle (*ḥarf*) describes what is neither noun nor verb” (Haywood 14). The fact that this simple division into parts of speech has stood the test of time and is in actual use to this day, not having been seriously challenged (Carter EALL Vol. II: 425)—besides marking the starting point for virtually all subsequent developments in grammar—demonstrates the enormous creative and formative potential of this precise categorization. Its importance is likewise seen in the issue of identifying loan words, which relies particularly on the ability to distinguish the characteristics of noun and verb respectively, and on recognizing the derivative relationship between them based on the words’ common roots (i.e. sequence of radical letters).

### 2.2.1.1 Nouns as elemental - distinguishing indigenous and non-indigenous modes of expression

To the Arabic noun is attributed the following features:

- (i) have declension, either full (*munṣarif*), partial (*ġayr munṣarif*) or invariable (*mabnī*);
- (ii) may be marked for definiteness or indefiniteness, by the prefix *al-* or *tanwīn* ‘nunation,’ respectively;
- (iii) have three numbers: singular, dual, and plural;
- (iv) have two genders, masculine or feminine. (Bernards EALL Vol. II: 424)

The declensional system (point (i)) in Arabic is in many ways unique and will be seen to form the primary means for distinguishing indigenous from foreign vocabulary in the language. As to definiteness (ii), al-Jawālīqī categorizes Arabicized nouns as having either Arabic or un-Arabic modes of expression as defined by their ability to take the definite article (*lām* of identification):

The Arabicized nouns, in [both] their conjugated and unchanged forms (*fī l-ṣarfī wa-tarkihī*) are of two kinds: The first is *not* counted among the un-Arabic modes of expression, namely the one to which is attached the *lām* of identification, such as *al-dībāj* and *al-dīwān*. The second *is* counted among the un-Arabic modes of expression, namely the one to which they never attached the *lām* of identification, such as *mūsā* and *ʿīsā*. (Al-Jawālīqī 53)

This formal differentiation would seem, however, only to identify the degree of the Arabicized words’ integration into the language, and is not known to have any practical consequences.

Compared with Persian, from which the majority of supposed loan words were thought to have originated (cf. Appendix 1) and whose nouns have only singular and plural forms, the existence of the dual (iii) in Arabic could potentially be of assistance for identifying loan words.

### 2.2.1.2 Verb as derivative of noun, indicative of process and time

The definition of the verb in Arabic evolved from being based on its morphological aspects, as seen in Sībawayhi's *Kitāb* and exemplified by his statement that it is a sum “of paradigms (*amṭila*) issued from nouns...” —more particularly from *maṣḍars* as a subclass of nouns; to that of focusing on its semantic aspects as indicated by process and time. Al-Zajjāj states with reference to Sībawayhi's definition:

[C]onventionally, according to the grammarians, the verb is what indicates a process and past or future time... This is what Sībawayhi meant by ‘as for the verb, it is a sum of paradigms issued from nouns depicting process and formed to indicate what has been, what will be but has not [yet] happened, and what is but has not been completed. (Hamzé EALL Vol. II: 90)

The standard definition, however, emerged with Ibn Sarrāj (d. 316/928) in his *Kitāb al-'Uṣūl* which states simply: “The verb is what indicates meaning and time, the past, present or future tense” (Hamzé 90).

#### 2.2.1.2.1 Importance of derivation in Arabic grammar

The fact of the existence in the language of a verb that was thought to have been derived from an existing noun (or sometimes vice versa) was considered to point to an indigenous origin; al-Jawālīqī refers to *al-muštaqq* “...the derivative [feature of Arabic grammar]” (al-Jawālīqī 51) and quotes Abū Bakr ibn as-Sarrāj (d. 316/928) from his epistle on derivation (*ištiqāq*) in the chapter *mā yajibu 'alā l-nāẓir fī l-'ištiqāq 'an yatawaqqāhu wa-yaḥtarisa minhū* ‘What the researcher of etymology must guard against’:

Among the things of which every cautious one should beware [of doing] is to make derivatives in Arabic of anything [that is actually] from a foreign language, for he would then be in the position of one who claims the bird to be born of the whale. (Al-Jawālīqī 51-2)



The application of the principle was, understandably, far from being an exact science. The word *qafṣ*, for instance, as listed in *al-Mu‘arrab* could go both ways:

*Al-qafṣ* is...genuine Arabic. It is from their utterance *qafaṣtu šay‘*. When you gathered it...And all things intertwined, thus *taqāfaṣa*... And some said: It is Arabicized Persian. Its original form (*‘aṣl*) is *kabastu*.<sup>2</sup> (p. 323)

Two further examples of exceptions that weaken the proposition by virtue of their being considered loan words possessing verbal derivatives are the cases of *al-‘urbān* and *al-kūs*. They illustrate as well the simple dynamics involved in the process of inventing new word-forms by means of derivation, and give a glimpse of the great potential that exists for expanding the language’s vocabulary:

Al-Farrā’ [said]: *al-‘urbān* and *al-‘urbūn*: A dialect [form] of *al-‘arbān* and *al-‘arbūn*. It is not said *al-rabūn*. It is a foreign word. The verb was derived (*ṣarrafa*) from it, so they said, “*‘arbantu fī l-šay‘*,” and “*‘a‘rabtu fī-hi*.” In the *ḥadīṭ* of ‘Umar: That he purchased the prison house with four thousand dirhams and “*a‘rabū fī-hā*”. Meaning: they loaned money...<sup>3</sup> (p. 280)

In the book of al-Mansūb [it] is related to al-Khalīl that *al-kūs* is a triangular stick [used by] carpenters with which they measure quadrangles (*tarbī‘*) of wood. It is a Persian word. Abū Hilāl said: The verb was derived from it, so they said *kāsa al-furs yakūs*... (p. 336)

A final example shows how the nonexistence of *ṣarf* ‘inflection’/‘derivation’ for the name *ṭālūt* ‘Saul’ as found in the Qur’ān was considered an indicator of foreignness:

*Ṭālūt*: A foreign name, God, the exalted, said: “And when Saul (*ṭālūt*) went forth with the hosts”<sup>4</sup> (Q 2: 249). The absence of its inflection (*ṣarf*) is evidence that it is foreign. Then if it were *fa‘alūt* from *al-ṭūl*, as *al-raḡabūt*, *al-rahabūt* and *al-tarabūt*: it would have been inflected... (p. 275)

<sup>2</sup> Wehr’s dictionary lists the noun *qafaṣ* ‘cage; pen; basket (made of palm fronds)’ (1994: 914)

<sup>3</sup> Wehr lists the form *‘arraba* ‘to give earnest money, give a handsel’ (1994: 702)

<sup>4</sup> Translation by Arberry

### 2.2.2 Abū l-Aswad's foundational role

Stories attribute Abū l-Aswad's motivation for teaching the new science to occasions at which he witnessed the potentially damaging results of mistaken language use, especially in Qur'ān readings but in everyday speech as well. Incorrect vowelling presented a notorious predicament, and he is thus said to have had the vowels of the text of the Qur'ān marked in order that they be correctly pronounced, and to have formulated the rules about *fā'il* and *maf'ūl* (roughly, 'subject and object'), *muḍāf* ('possessed'), and of *naṣb*, *raf'*, *jarr* and *jazm* ('accusative', 'nominative', 'genitive' and 'jussive' respectively). It is reported that he was a *qārī'* (Qur'ānic reader), which would further vindicate his readiness for the task (Haywood 11-12).

The corruption of Arabic speech in the newly conquered territories is cited as well by Ibn Khaldūn (d. 808/1406) as the impulse for the development of grammar, the function of which was to teach new users of the language how to speak it correctly, thereby reducing the chances for further corruption (Versteegh 2006: 3).

If the activity seen here of discerning the identifying characteristics of the language was itself any measure, the subsequent step of discovering its underlying structures was not far distant; and the third step, the establishment of an exclusivist mindset about the language with its concomitant notion of foreignness, could be considered to be brewing. Unfortunately, there is little information to be gleaned about developments actually taking place in the earliest stages of Arabic linguistic science beyond scattered anecdotes of the type mentioned, and we must instead rely on works of collected biography such as the one written by the well-known Andalusian lexicographer Abū Bakr Muḥammad ibn al Ḥusayn al-Zubaydī (d. 379 A.H.), who collected biographies of philologists in his *Ṭabaqāt al-naḥwīyīn wa-l-luġāwīyīn* (Haywood 16). He thus introduces Abū l-Aswad:

He was the first to establish (the science of) the Arabic language, to lay down its methods, and to establish its rules, and that was (at a time when) the speech of the Arabs became disturbed, and people high and low came to make mistakes... (Haywood 12)

*Al-Mu‘arrab* also contains reference to the general “confusion” among the Arabs that arose as a result of the interference of loan words, as seen in this quote attributed to Abū ‘Umar al-Jarmī:

The Arabs may have confused [certain] foreign [terms] (*al-’a‘jamī*) when they brought them into their [own] language (*luġa*). On the authority of Abū l-Mahdī, he said: The Arabs have caused confusion in the matter [of foreign vocabulary (*al-’a‘jamīya*)] and speak about it in an inconsistent manner since it is not [an integral] part of their [own] speech. (p. 56-7)

### 2.2.3 Twofold specialization merging in the subject of derivation

The first steps in the analysis and systematization of the language, following the codification of the holy book, had thus been taken. These studies later split into the twin sciences of grammar (*naḥw*) and lexicography or philology (*luġa*). Despite their strong interdependence, the function of each was clearly defined: The lexicographer (*luġawī*) was charged with ensuring that the pure speech of the Arabs be recorded and handed down in vocabularies and dictionaries. The grammarian had to show how this material was used in connected speech, stating the relevant rules that were arrived at through analysis and synthesis. The two overlapped in the subject of derivation (Haywood 17-18), which was as much the keystone in distinguishing Arabic from un-Arabic modes of expression as *qiyās*—upon which the notion of derivation was built—has been found to represent a starting point for basic epistemological procedures in all the cultures of the world (Maroti EALL IV: 11). From this perspective, it could perhaps be surmised that progress in the matter of distinguishing foreign vocabulary would represent some kind of indicator of the state of development of the science of linguistics itself.

The expanded use and resulting refinement of the implements of derivation and *qiyās* as applied by al-Khalīl in his lexicographical work *Kitāb al-‘Ayn* and by Sībawayhi in his grammar *Kitāb* respectively can, correspondingly, together be seen to have occasioned comparable advances in the field of distinguishing loan words in the language. The interest of both linguists in the subject is clear from their focus on identifying criteria for the morphological structure of Arabic words, evident as well in Sībawayhi’s use of the expression

*laysa fī kalām al-‘arab* ‘not [found] in the Arabs’ speech’ in reference to certain structural patterns (Baalbaki 2008: 225).

It may be noted that the ‘*Ayn* and the *Kitāb* would establish the essential forms the disciplines of lexicography and grammar were to take in the subsequent 200 and 1000 years respectively (Haywood 19).

## 2.2.4 Concerns of the early linguists

Little is known about the linguists that lived between Abū l-Aswad and al-Khalīl but their names, such as that of Yaḥyā ibn Ma‘mar and ‘Īsā ibn ‘Umar al-Thaqafī (d. 149/766), the titles of certain of their works, and the fact that their interests coincided with religious studies. ‘Īsā is said to have further developed the grammar of *fā’il* and *maf’ūl* as taught by Abū l-Aswad. Their principal preoccupation was with what today is called *qirā’āt* ‘readings’ of the Qur’ān—the deficient, unvowelled early text existing in many variant readings and requiring collation so as to determine their acceptability according to the reliability of the transmitters and their conformity with the linguistic usage of the Bedouins of Central Arabia in pre- and early-Islamic times (Versteegh 1997: 25-26; Bohas et.al. 1990: 2). The entry *mīkā’il* ‘Michael’ in *al-Mu‘arrab* exemplifies the nature of the discussions surrounding variant *qirā’āt*:

*mīkā’il*, Ibn ‘Abbās said: *jabrā’il* and *mīkā’il*: *jabr*: Servant, as in your speech: ‘Abd Allāh and ‘Abd al-Raḥmān. It meant that *’il* was the name of God, the exalted, and the name of the angel *jabr* and *mīkā*, so they were related to God, the exalted. The exegetes have not differed on this. [But] the [Qur’ān] reciters (*qurrā’*) differed in their readings (*qirā’a*): So some of them read *mīkā’il*. And some *mīkāḷ*. And others *mīkā’il*. And Ibn Muḥayṣin recited *mīka’il*. Like *mīka’il*. Al-Ḥarbī said: Abū ‘Umar informed me from al-Kisā’ī saying: *jibrīl* and *mīkā’il* are names that the Arabs could not have known, so when they came, they Arabicized them. (p. 375)

Another central concern among early scholars was the collection and assessment of ancient poetry, which would enlarge immeasurably the canonical corpus upon which philological study was based. The problems involved in priming this variegated material for use as grounds for valid deductions in linguistic questions were much more complex than they had

been with the relatively consistent Qur'ānic textual variants; the subject of poetry was very different from that of the holy book, being concerned primarily with worldly affairs and admitting a greater measure of technical vocabulary, rare words, involved constructions and tribal idioms. An example of a technical word having idiomatic forms and attested in poetry is the word *al-nawraj* 'thresher' with its possibly related form *al-narja*, as listed in *al-Mu'arrab*:

Al-Layth: *al-nawraj* and *al-nayraj* are two dialect [forms]. The people of Yemen say *nūraj*. It is what food is crushed with, of iron or wood. The poet said: "... *'kamā yaṣirru l-nawraj*..."

'Ammār ibn al-Bawlānīya said: "...*hādā l-ladī yajrī 'alayhi l-nawārij*." (p. 383-4)

Al-Azharī related from Ibn Durayd: *al-narja*: Piece of wood that overturns the earth. And in the rare inflections (*nawādir al-'a'rāb*): *al-nawraj*: mirage. And *al-nawraj*: lane of the plowman. Al-Layth said: *al-nayraj*: a spell,<sup>5</sup> as magic..., it is certainly an illusion (*tašbīh*) and deception. All of this is imported, since the *nūn* and *rā'* are not combined in one word in the Arabs' speech. (p. 385)

The transmitters of poetry were also less careful than the Qur'ānic readers to preserve the exact wording of the selections handed down. There had been a tendency among them—typical of oral traditions—to modify the old poems as they saw fit, and even to interpolate, at times, their own lines in the course of recitation.

Certain branches of philological science (*'ulūm al-'arabīya* 'Arabic sciences') gained particular importance in response to such complications of transmission: lexicography (*'ilm al-luġa*), 'rare expressions' (*ġarīb*), metrics (*'ilm al-'arūd*), and, to a certain extent, knowledge of legendary battles and tribal wars of the ancient Arabs (*'ayyām al-'arab*) as well as their genealogies (*'ilm al-'ansāb*). These fields would serve as *ad hoc* frameworks for understanding the profuse, creative allusions that permeated such works; the latter two were necessary for understanding recondite references to tribal rivalries and alliances and the complex relationships underlying them (Bohas et.al. 2-3).

Since the conceptual framework necessary for distinguishing foreign elements in the language had not yet been developed, the absence of reference to the subject of loan words

---

<sup>5</sup> Shākir, editor of al-Jawālīqī's *al-Mu'arrab* (1969), defines the word used, *aḥḥad*, as *ruqya* 'spell, charm, magic,' as listed in Wehr (1994: 411)

should not come as a surprise; at this early stage it could perhaps be considered as implicit, belonging under the rubric *'ilm al-luġa*.

### 2.2.5 The earliest word lists

The philological work before al-Khalīl had primarily been conveyed orally, and whatever written work there was in lexicography and grammar would have been subsumed in the *Kitāb al-Ayn* and Sībawayhi's *Kitāb* respectively. Any works that those lesser known pioneers are said to have written have not survived, but it is determined that *ġarīb*, especially as found in the Qur'ān, was among their specialties. Haywood (1965: 41) labels this particular interest a 'cult' phenomenon, and concludes: "The parade of this sort of erudition never ceased to be a feature of Arabic lexicography" (Haywood 17-19). The evidence of the numerous titles of specialized vocabulary lists found in the *Fihrist* of Ibn al-Nadīm (d. c. 385/995) clearly corroborates this view.

The organization of the entries in these early specialized lists of strange and unusual words in the Qur'ān simply followed the order of occurrence in the latter text or were arranged semantically (Versteegh 1997: 27). They were thus not, strictly speaking, dictionaries, but more like short monographs; yet their focus on the *meaning* of words represented the starting point for later, more systematic developments in Arabic lexicography. A transitional phase could perhaps be seen in works treating circumstantial fields like word etymology and special features such as homonyms and words with opposite meanings (*'aḍḍād*), which in turn engendered vocabularies focused on one semantic domain, such as the horse, the camel, a plant variety or human anatomy (Versteegh 1997: 21-22).

A look at a typical example of an entry in one of these earliest word lists on the terminology of palm trees and grape vines, ascribed to al-'Aṣmā'ī (d. 831), is highly instructive, revealing the focus of the lexicographers' interest:

Young palm trees are called *jaṭīṭ*; (they are the first that sprout from the mother tree.) They are also called *wadī*, *hirā'*, and *fasīl* "base, ignoble, offset". When the offset is still attached to the stump and has not yet rooted it is called *ḥasīs al-naḥl* "the vile part of the palm tree". The Bedouin call it *rākib* "rider". When the young sprout is torn from the mother tree together with the stump of its branch, it is said to be *mun'ala* "shod". When it is planted, they dig a well for it and plant it there, then they fill it up

all around with slime from the river and dung. This well is called *al-faqīr* “the poor”. The expression is: *faqqarnā li-l-wadīya* “we dug a well for the offset”, verbal noun *tafqīr*. Another name for a young palm tree is ‘*aṣa*’. (Versteegh 1997: 26)

The emphasis in the former entry is on Bedouin lore, their expressions, idioms, customs and practices, and such treatises also naturally covered proverbs and poetry. The compilers were especially attracted by rare terminology, a predicament in keeping with the limited scope of their work, focused as it was on words which were falling into disuse, existed only in the pre-Islamic poetry or in specific tribal dialects (Versteegh 1997: 26).

An entry on the same subject from al-Jawālīqī’s eleventh century *al-Mu‘arrab* reveals the linguists’ continued interest in such apparently inconsequential matters:

*Al-birzīn*: Arabicized Persian. It is the container (*inā*’) of the peel/skin (*qiṣr*) of the inflorescence of the palm tree (*ṭal*’). The Arabs have given word to it. The Baṣrans call it *al-taltala*. ‘Abd al-Raḥmān explained it thus from his paternal uncle...  
(Al-Jawālīqī 117)

## 2.2.6 First major systematizations of the language

These initial attempts at patching the uncharted territory of the seemingly unbounded wealth of the Arabic lexis opened the way for the genius of al-Khalīl to conceive of cataloguing the entire inventory of the language according to a systematic arrangement of the words’ letters by their root consonants (including *wāw* and *yā*’). This far-reaching innovation was soon followed by a comparable feat in the related field of grammar, in the form of Sībawayhi’s monolithic *Kitāb*, which organised and harmonized what was known by then about the structure of the language while at the same time creating the first exemplar of a ‘book’ in the full sense of the term, which implied a wholeness in structure and content including cross-references (Versteegh 1997: 40). The Qur’ān itself was certainly a *kitāb* ‘book’ by its own confession, the word occurring 261 times therein (Madigan EQ Vol. I: 242) in a wide variety of contexts; yet the several levels of meaning the word takes on in the Qur’ān sets it clearly apart from the simpler meaning normally connected with it.<sup>6</sup>

---

<sup>6</sup> Examples of Qur’anic meaning range from ostensibly corporeal references, as in the verse

A system had already been devised for representing the special feature of Semitic languages, of which Arabic is a very typical example: viz. that the configuration of consonants making up the radicals of a word defines its semantic value. The letters of these radicals, usually three in number, were represented by the sequence *f*, ‘ (‘ayn) and *l* in that order, a system whose inherent potential for describing the morphological, phonetic and semantic features of words was progressively realized by al-Khalīl and Sībawayhi, and which became definitively synopsized in the latter’s theory (cf. section 2.5.1.4.1.1 below).

## 2.2.7 Relationship of the linguistic disciplines to Qur’ān exegesis

The sum of these rapid advances in lexicography and grammar provided the technical and conceptual apparatus needed to explain Qur’ānic usage; yet it was Qur’ān exegesis itself that remained the decisive occupation of scholars in Islam during the early centuries, even if this fact often was only indirectly acknowledged.<sup>7</sup> The field would develop through the close dialectic relationship it maintained with those subsidiary disciplines, all of which can be regarded as elements of a single great enterprise directed toward the clarification of God’s intention through analysis and exposition of the sacred text and message. The exegetes were thus dependent on receiving a thorough training in grammar to carry out their work, which they would put to use in compiling commentaries specializing in fields such as textual variants, grammatical and syntactic analysis, analysis of narratives in the text, and law (Versteegh 1997: 22), as well as occasionally pronouncing upon questions of etymology and foreign vocabulary. The early philologist Abū ‘Ubayda (b. 110/728), for instance, although a staunch believer in the pure Arabness of all words found in the holy book, is seen to address

---

*yawma nad‘ū kulla ‘unāsin bi-imāmihim faman ‘ūtiya kitābahu bi-yamīnihi fa-‘ulā’ika yaqra’ūna kitābahum wa-lā yuḏlamūna fatīlan* (Q 17:71) ‘On the day when We shall call all men with their record, and whoso is given his book in his right hand – those shall read their book, and they shall not be wronged a single date-thread.’ (Arberry)

to clearly allegorical ones: *yamḥū llāhu mā yašā’ wa-yuḏbitu wa-‘indahu ‘ummu l-kitāb* (Q 13:39)

‘God blots out, and He establishes whatsoever He will; and with Him is the Essence of the Book’. (Arberry)

<sup>7</sup> “Despite intellectual controversies about the role of Reason and Revelation in Islam, the paramount significance of Revelation was never minimized. Throughout the ages philosophers tried hard to bring harmony between Reason and Revelation... No Muslim Philosopher did ever deny the Quran as the basis of Islamic metaphysics.” (Nadvi 1997: 139)

“Al-Shāfi’i’s (d. 204/820) position is one that concurs with his legal reasoning: the knowledge of the Arabs in language is a part of “tradition” which must form the basis of Muslim society. The study of language, like the use of reason in law, has its place, but it must always come second in significance and authority to traditional knowledge.” (Rippin, EQ vol. III: 230)

Kopf (1956: 33-34), on the other hand, considers the view held by “the Arabic tradition as well as by modern scholars,” that “Arabic philology had come into existence as a means for the interpretation and correct reading of the Holy Scriptures of Islam” to be “wrong or at least one sided.”



the specific case of the Qur'ānic word *ġassaq* (Q 78:25) that seems to have been presented to him, as related in *al-Mu'arrab*:

Ibn Qutayba said: Abū 'Ubayda did not believe that there was anything in the Qur'ān that was not Arabic. He had said: It is a coincidence that has occurred between the two languages. Others beside him claimed that *al-ġassāq* [is] malodorous cold in the Turkish language. It is said: [Its pattern] is *fa'āl* from *ġasaqa yaġsiqu*, this is actually Arabic. It has also been read with a weakened consonant (without *tašdīd*), and it is like [the pattern of] *'aḏāb* and *nakāl*. It has been said regarding its meaning: it is severe cold, burning cold. It is what flows of pus from the skin of the people of fire ('hell'). (Al-Jawālīqī 283)

## 2.3 Al-Khalīl ibn Aḥmad and his major work, *Kitāb al-'Ayn*

As already mentioned, the work of al-Khalīl became the model upon which the fledgling science of lexicography developed, thereby influencing the entire subsequent course of the discipline. He was an Arab, having been born in Oman in 99/718 but moving to Basra at an early age. As was the case with most of the linguists before and during his time, al-Khalīl's scholarly work began from Qur'ān exegesis and knowledge of the *sunna* (manner or deeds of Muḥammad), and only later expanded into the sciences of lexicography, grammar, the *ṣarī'a*, mathematics and music, besides occasional excursions into poetry. He is credited with inventing not only the Arabic lexicographical discipline as we know it, but also musicology and metrics, and is said to have a hand in the reform of Arabic script. He was also the main teacher of the illustrious Sībawayhi. He died in 175/791 (other dates given are 159/776 or 170/786) (Haywood 21; Sellheim EI Vol. IX 962; Versteegh 2001: 62).

Al-Khalīl's diverse interests all show his unique ability to discern sounds and their interrelationships. His areas of study such as the phonotactic rules of Arabic morphological roots, grasp of phonological structure, analysis of poetic meter, and interest in music, together reflect a flair for identifying sound patterns and relationships across disciplines. His interdisciplinary undertakings are evidence of a context-independent cognitive process, which might be compared to the faculty of reasoning by analogy so central to Arabic linguistic thought (Ryding 1998: 8) (cf. section 2.5.1.3.2 below).

### 2.3.1 *Kitāb al-‘Ayn*

#### 2.3.1.1 Historical assessment, scope and formative influences

The most famous work attributed to al-Khalīl is the *Kitāb al-‘Ayn*, known as the first dictionary of the Arabic language. There is, however, no oral tradition through a chain of successive scholars establishing definitely al-Khalīl’s authorship of the book, although he is universally acknowledged to have initiated it. The most common view is that someone else, possibly his one-time pupil al-Layth ibn Naṣīr ibn Ṣayyān, finished the uncompleted work on al-Khalīl’s model. Referring to its many errors and shortcomings, al-Zubaydī (d. 379/989) in his abridgment *Muḥtaṣar kitāb al-‘ayn* gives a fair assessment that takes into consideration al-Khalīl’s recognition and towering stature as an exacting, eclectic scholarly genius: “Al-Khalīl laid down the lines of the book, and arranged its division into chapters; but others who were unreliable filled out this skeleton” (Haywood 53). Tributes to his life’s work draw attention to a number of significant indicators of his achievements:

Nothing can detract from al-Khalīl’s genius. To have conceived the idea of a comprehensive Arabic dictionary, even with the help of the ideas of other men and peoples, and to have started writing it, is achievement enough for any eighth-century Arab. After all, no-one denies al-Khalīl credit for codifying Arabic prosody, even though his book on the subject is not extant. (Haywood 27)

The Arabic language in its entirety (both prose and verse) became the object of al-Khalīl’s extensive research. The durability of his achievement is truly astounding since Arabic grammar, lexicography and prosody are now essentially where he left them twelve hundred years ago. (Shahīd 1998: ix)

Al-Khalīl was a genuine innovator, as his works on metrics, grammatical *‘illa* (causation), and phonetics clearly demonstrate. ...his concern for establishing a solid base for grammatical study can best be shown in the terminology he introduced... [H]is attempt at discovering the boundaries of linguistic material...is the most manifest proof of his linguistic insight. (Baalbaki 1998: 46)

Al-Khalīl's piety has been regarded as a determining factor for his linguistic genius, an attribution which carried not little importance in Muslim society and can be regarded as the ultimate seal of endorsement in the Islamic sciences. It is indeed doubtful whether attention would have been given to any scholarly work that was felt to be detached from the strong religious current of the time, as demonstrated by the following assessments:

The notion of authority being derived from piety is a central proposition... A remarkable feature of Islamic science, unmistakably demonstrated by al-Khalīl, is that a scientific truth acquires its acceptability not from the methods or principles that produced it, but from the personal aura of the individual who enunciated it. (Carter 1998: 32)

...the strong religious sentiment he was possessed of may be conceived as the animating and unifying force behind his endeavors on behalf of Arabic. (Shahīd 1998: x)

Al-Khalīl is...a paradigm for practically the whole range of activities proper to the Muslim intellectual: He is not only pious far beyond the conventional and completely familiar with the whole Islamic syllabus, but he is also a creative thinker...and a master of Arabic, both linguistically and stylistically combining the attributes of the sage and courtier within the person of a scholar. (Carter 1998: 33)

### **2.3.1.2 Generative conception**

The "Fihrist" quotes al-Khalīl addressing al-Layth: "If someone made a plan, and wrote the letters alif, bā', tā', thā, and so on, he would then include all the language of the Arabs... He should arrange it under biliterals, tri-literals, quadriliterals, and quinquiliteral roots. There is no speech known to the Arabs with more than that" (Haywood 24-25). The few manuscripts that today remain of the 'Ayn, as much as the seminal influence it had on the entire lexicographical tradition, confirm his achievement of this monumental task.

Implicit in the above quotation is the pregnant suggestion that all originally Arabic words have been coined from one or another of these roots. The intention was thus primarily to identify all root combinations from which words in actual use had been derived and

secondarily to document the latter, usually excluding the most common forms of the root which were expected to be known to the reader. This very specific purpose of al-Khalīl's dictionary identifies it as a strictly scholarly production, a genre that ruled virtually supreme in the realm of literary reproductions in the early formative period of Islamic science since only a fraction of the ruling Arab population was literate. Neither can it be expected to conform to any paradigm of dictionary production known today, simply because such paradigms could only be developed over time.

There can probably be no clearer statement than the above final emphatic words of al-Khalīl, "there is no speech known to the Arabs with more than that," of the nature of the project he initiated viz. to distinguish clearly Arabic from non-Arabic formation.

### 2.3.1.3 Organization of entries

The corpus of roots listed in the *ʿAyn* might have been expected to be ordered according to the usual method of the alphabet, starting with *alif*; but al-Khalīl's rationalizing mind, fired by his zeal to uncover the natural order existing in every phenomenon, as much as his specialization in the phonetic aspects of the language, decided him on a different arrangement: the first letter of the alphabet, *alif*, did not have any known characteristics that would qualify it for such a privileged position; for one, it was a weak consonant; second, it had the special, unenviable status in Arabic phonetic theory of being an unpronounced and thus abstract phonological element; third, long vowels were not acknowledged by the Arabic grammarians, rather, they regarded them as combinations of (short) vowel and glide (/ʾ/, /w/, /y/), i.e. /aʾ/, /uw/ and /iy/, which are written conventionally as [ā], [ū] and [ī]. The two distinguishing features of *alif* as compared with the /w/ and /y/ are thus, on the one hand, that it is usually not realized on the phonetic level, appearing as a dummy letter; and on the other, that when it is actually pronounced, it is realized phonetically as a glottal stop /ʾ/ or assimilates to an adjacent /w/ or /y/. The grammarians' analysis was not based on the Arabic script, which represents the long vowels in the simple form of the letters *alif*, /w/ and /y/, but on the structure of the language as it was pronounced: by ear and by way of phonetic transcription, identical word patterns could be identified in seemingly dissimilar words, such as *ṣufr* 'yellow [plural]' and *sūd* 'black [plural]', phonetically transcribed as /ṣufr/ and /suwd/ respectively (Versteegh 1997: 27).

Al-Khalīl thus went in search of a suitable alternative for initial letter of the alphabet basing himself, once again, on phonetic criteria. The choice fell on the simple ordering principle of the letters' articulation points. The first letter was set to be the one pronounced deepest in the throat, namely *ʾayn*, and the subsequent ones proceeded in their natural order to the lips, from gutturals to velars and so forth, ending with *bā* (Haywood 38-39). Among the possible influences that have been suggested for this phonetic-permutative ordering method is India (Sellheim 963), but inquiries in this regard have so far been inconclusive. The system proved in the long run to be impractical compared with an alphabetical arrangement, by which it was eventually superseded. Further to this unusual arrangement, his dictionary ordered the words anagrammatically in groups around the permutation of each set of radicals, a system which in time also came to be abandoned (Haywood 38-39).

The underlying supposition apparent in al-Khalīl and other linguists' approach—that of an inherent order existing in and permeating all aspects of the Arabic language—was eventually followed through to its logical conclusion in linguistics, though not always with a successful outcome. The system of the permutation of the radicals of words as the organizing principle as used by al-Khalīl came to be applied to other areas of the language as a factor for distinguishing Arabness, notably in semantics. The famous philologist Ibn Fāris (d. 390/1000), who showed an equal interest in jurisprudence (*fiqh*) and lexicography, wrote a book entitled *Ṣāhibī fī fiqh al-luġa* (*ṣāhib* referring to his patron Ṣāhib ibn ʿAbbād) which was a kind of encyclopedia of Arabic lexicography that made the interesting connection of the term *fiqh* with language, thereby making explicit the intuitive supposition of the linguists. His project was the establishment of a closer link between the two disciplines, and he drove to integrate philological studies in the schooling of jurisconsults. Haywood (1965: 100) argues that by making the connection of *fiqh* with language he was suggesting that the latter was as scientific and logically organized as the former. This idea is supported by the scientific and logical approach found in Ibn Fāris' two dictionary works as well as in his treatise *Kitāb al-Ṭalāṭa* 'The Book of Three' in which he collected trilateral roots whose three letters were thought to hide a basic meaning irrespective of their order of combination. His contemporary Ibn Jinnī (d. 392/1002) was occupied with the same subject. The idea has, however, been found to have very limited applicability and it is not generally recognized (Haywood 98-100).

### 2.3.1.4 Constitutive theory and basic intention

The theory of language found in the *ʿAyn* must have existed in some rudimentary form prior to the book’s conception as developed by his predecessors, but it could only start to take shape after the adducing of a sufficient number of observed items and cases that had been compared, analysed and categorised (as had been done in the early word lists), thus giving the impulse to systematization (Sellheim 962). The fact that his magnificent pupil Sībawayhi quotes him in support of his statements in no less than 608 out of the almost 900 such scholarly references in the *Kitāb* is ample testimony to al-Khalīl’s capacity for creating and organizing representative samples of the language (Haywood 22; Baalbaki 2008: 16). The purpose of the work had been to distinguish between roots in use – *mustaʿmal* ‘used’ – and those not in use – *muhmal* ‘neglected’, i.e. not found in Arabic. As to the rationale behind the selection of words for inclusion in the *ʿAyn*, Versteegh explains: “When words derived from a root are mentioned, this solely serves the purpose of showing that the root actually exists in the language” (Versteegh 1997: 29).

As was the case with the earliest word lists, the contents cater mainly to the less common words, many of which in practice would already have gone out of use but whose existence, as purveyed in verses from poetry or the Qur’ān, served to validate the root combination in question. The confirmation of such existence would then open the door—at least in theory—for derivative forms of the root in question to be conjured when needed. It can be noted that the practice of making analogical creations from existing roots was used extensively in the nineteenth and twentieth centuries to express modern ideas, concepts and phenomena that were being introduced into the Arabic language originally as loan words from the enterprising West.<sup>8</sup>

Such an understanding of the generative potential of the Arabic root system would have been firmly, if only implicitly, established in linguistics by the time of al-Khalīl; it only remained for him to articulate the notion, in support of which his *ʿAyn* provided as much circumstantial evidence, in the form of recorded word items in lemmata, as could be mustered at the time.

---

<sup>8</sup> For a survey of practices in the Muslim world with regard to the adoption of foreign vocabulary in Arabic in modern times, refer to Stetkevych (1970).

## 2.4 Subsequent developments in lexicography

Lexicographers after al-Khalīl took the *ʿAyn* as their model and simply expanded upon the entries, introduced certain technical, grammatical and etymological labels to classify them, and reorganized their order of arrangement to suit their particular needs and concerns. The field was cumulative in the sense that dictionaries were continually expanded with new and rare vocabulary and meanings that were being discovered, a fact that also opened for criticism of earlier works. The lexicographers' initial bias toward and prioritizing of rare and curious terminology for inclusion in their works was gradually superseded by a desire to include the entire lexicon. What Calder (1993: 133) has termed the "acquisitive" tradition of Islamic exegesis, which would naturally extend to lexicography and etymology as well, implied that whatever judgments had been handed down, up to about the eighth/fourteenth century, had to be accepted uncritically as an inalienable element of the contemporary and future corpus of knowledge, not be subjected to reform or, much less, elimination (Rippin EQ Vol. III: 232).<sup>9</sup> Kopf points to the same circumstance:

...the fact that, in indigenous Arabic dictionaries, we find so many unwarranted explanations of Qur'ānic and Ḥadīth words originating in traditional and even sectarian exegesis. Lexicographers could not easily dismiss even doubtlessly mistaken interpretations or replace them by better ones which they might have been able to establish with the professional means at their disposal... Arabic lexicography, soon after its inception, gradually developed into a traditionary discipline to which the principles of the science of Ḥadīth came to be applied. Religious factors are said to have been brought into play, when it was intended to establish the reliability of the "transmitters" of the language. Al-Suyūti states that the lexicological traditions of libertines were rejected... Tradition has it that religious considerations also influenced the choice of poets whose verses were to be taken as *šawāhid*. (Kopf 1956: 38-39)

---

<sup>9</sup> For a survey of historical developments in the Arabic lexicographical tradition up until modern times, refer to Versteegh 1997: 29-35.

## 2.5 Sībawayhi

*Sībawayhi* (d. approx. 180/796), whose full name, though it is not in actual use, is said to have been Abū Bishr ‘Amr b. ‘Uthmān b. Qanbar, and who was a *mawla* of Banū Ḥārith b. Ka‘b Sībawayhi, is the peerless grammarian and founder of the Arabic grammatical tradition as we know it. His birthplace is supposed to have been al-Bayḍā’, Shirāz, whence (the province of Fars) he returned, dying at the young age of between 32 and 40 years. He came to Baṣra to study *ḥadīth* and jurisprudence (*fiqh*), but his interest soon shifted under the influence of his various teachers—among them Ḥammād b. Salama (d. 167/784) and Yūnus b. Ḥabīb (d. 182/798) but more notably al-Khalīl b. Aḥmad—to the budding field of grammar. His unique precocity and insight set him off from his contemporaries as much academically as it did socially, which is a probable reason for the extreme dearth of identifying anecdotes that have reached us about his life. He seems to have kept aloof from the insular controversies that surrounded the scientific milieu of his time, instead focusing his extraordinary intellectual powers on devising a comprehensive, flexible and cohesive system for understanding the structural characteristics of the Arabic language, and on discovering the generative principles underpinning them (Carter 1973: 146-7).

### 2.5.1 The *Kitāb*

The appearance of Sībawayhi’s grammatical *oeuvre*, posthumously given the simple honorary title of *al-Kitāb* or *Kitāb Sībawayhi*, also known under the title *Kitāb fī al-naḥw*, constitutes a signal event in the Islamic scientific enterprise, virtually unsurpassed to this day for the originality of a linguistic system developed from the vastest of research, and for the pervasiveness of its influence. He was diligent, as well, in citing the sources for his material and faithful in propagating the lessons and technical terminology of his predecessors, expanding and revising it in the process of forming his general theory. The complete disappearance of all works on grammar known to have existed prior to the *Kitāb* is a telling sign that their subject matter had been subsumed in Sībawayhi’s work, obviating thereby the need for their reproduction and propagation. In noting and describing a vast breadth of attested usage in Arabic as conveyed by his precursors and contemporaries, it was Sībawayhi’s privilege to combine the data he collected into a meaningful whole and discuss it



in terms of the social, psychological, and strategic decisions the individual speaker makes in communicating ideas, in consonance with the reciprocal response to these choices by the listener (Carter EI Vol. IX: 527).

### 2.5.1.1 Contents and influences

The work consists of seven introductory chapters followed by chapters on syntax, morphology and phonology, forming a large yet consistent and systematically ordered book with internal cross-references. The introductory subjects are the three parts of speech, the sets of vowels and inflections, certain internal hierarchies, a basic subject-predicate arrangement, various lexical, semantic and phonological phenomena including synonymy, polysemy, elision and substitution, an enumeration of formal and semantic criteria, and an examination of unconventional forms occurring only in poetry (Carter EI Vol. IX: 526).

Sībawayhi's use of technical terminology is very consistent, conforming closely to that of his predecessors, while at the same time departing significantly from it in many respects. An example will demonstrate the nature of developments he fostered in the area of the technical apparatus of grammar: In his structuring the declensional scheme, Sībawayhi for the first time makes the distinction between declensional and other vowels: those produced by 'āmil 'operant'/'governor' and those not affected by syntax. For the former he uses the terms *raf'*, *naṣb*, *jarr* and *jazm*, and for the latter *ḍamm*, *fath*, *kasr*, *waqf*. It also happens that the same term features with what appears to us as different meanings in respectively its general and purely technical sense, such as that of *muḍā'ir* and its derivatives, 'āmil, fā'il, binā', *tamakkun* etc. The term *fi'l*, for instance, is seen to signify both 'action' and 'verb', *ḥāl* both 'conditional and 'circumstantial accusative', and *zarf* both 'circumstance' and 'adverb' (Baalbaki 2008: 32-33).

As to the influence of al-Khalīl in the production of the work: he is Sībawayhi's primary reference source, being cited 608 times out of the almost 900 such references in the *Kitāb*, besides being known to be the originator of much unattributed material as well in the work (Baalbaki 2008: 16-17). He thus proves to be greatly dependent upon the insight and expertise of his main teacher, yet his vast and enduring theoretical and empirical contributions to Arabic linguistics, not the least of which is to be seen in the unified grammatical theory and methodology that is presented in the *Kitāb*, demonstrates clearly his originality.

### 2.5.1.2 Sources for his work

As mentioned, the research activity of the early linguists was closely linked to Qur'ān exegesis and the analysis of Qur'ānic usage. The non-Qur'ānic data, which constituted the bulk of available sources, was used for this same purpose of interpreting the holy text, and it would probably have required a degree of courage from a Muslim linguist to shift the focus of his interest away from the Qur'ān, representing the known system *par excellence* as it did, so as to be able to discover another related yet distinct system in the breadth of the Arabic language itself. This was the task that Sībawayhi had set himself. This very ambitious plan at times compels him to walk a tight rope when dealing with the *qirā'āt* 'readings' of the Qur'ān, balancing his wont for giving value judgments on variants, on the one hand, and of avoiding theological issues related to them, on the other, a dilemma he is not always able to resolve. An example is his treatment of *mā* as negator of nominal sentences. The Ḥijāzīs have the predicate in the accusative form (e.g. *mā Zaydun munṭaliqan*) whereas the Tamīmīs use the nominative (*mā Zaydun munṭaliqun*), and Sībawayhi has no qualms about expressing his preference for the latter as linguistically more 'correct' in relation to the system of *qiyās* he used to analyze it; this despite the fact that Qur'ānic usage conforms consistently to the former, as in *mā hādā bašaran* 'No mortal is this' (Q 12:31), where the final *alif* of the text allows for this reading alone (Baalbaki 2008: 36-37). The speech of the Arabs (prose) and poetry, on the other hand, carried no religious implications, and would have afforded Sībawayhi freer scope for the exercise of his judgment.

#### 2.5.1.2.1 *Samā'* 'attested data'

The importance of Sībawayhi's use of the concept *samā'*, which represents the sources from which his data is derived, lies in the centrality of this practice for the whole of the grammatical enterprise. It embodies, besides expressions derived from the root *s-m-* 'such as *samā'* and *sam*', terms such as *ra'aynā l-'arab*, *sa'alnā...*, *yaqūlūna*, *yu'ḥad min al-'arab*, *min al-'arab man...*, *min kalām al-'arab*, *ḥaddaṭanā man yūṭaq bi-hi*, *balāḡanī 'an al-'arab al-mawṭūq bi-him* 'annahum yaqūlūna, *za'ama lī ba'd al-'arab*, *sa'alnā l-'arab fa-wajadnāhum yuwāfiqūnahu*, etc. (a list of corresponding expressions occurring in *al-Mu'arrab* is given in Appendix 2). In the same category is his reporting of *riwāyas* on the authority of recognized sources, including his teachers. The body of transmitted data used in

the *Kitāb* has been referred to as *naql* ‘transmission’, in contrast with the material invented by the grammarians themselves through *qiyās* which would not have been supported by actual usage.

The attested material in the *Kitāb* can be divided into four categories: the Qur’ān, the Islamic traditions (*ḥadīṭ*), the speech of the Bedouin (including proverbs, speech patterns and idiomatic expressions) and poetry. Sībawayhi tailors his methodological approach to the unique characteristics of each of these. As to *ḥadīṭ*, he quotes it rarely, as do most other grammarians, supposedly due to problems related to its transmission: being largely non-verbatim, it was generally unreliable. Qur’ān posed the twofold problem of the existence of variant *qirā’āt*, about some of which Sībawayhi would have cause to be sceptical; and the theological issues already mentioned (cf. section 3.5.1.1 above) (Baalbaki 2008: 35-36).

There was during Sībawayhi’s time and in the following decades, a vogue for collecting linguistic attestations (*šawāhid* pl. of *šāhid*) of usage from Bedouin informants, and a huge corpus of material was thus available for linguists to draw upon in their research (Baalbaki 2008: 26). The distribution of his sources has been estimated to 1050 lines of poetry (*šī’r*), 447 Qur’ānic verse, 350 speech patterns or idiomatic expressions (*kalām*) and 41 proverbs. He also invented a number of examples (*amṭila*, pl. of *miṭāl*) of phrases that could occur in theory for the purpose of illustrating syntactical relationships; although it is not always clear whether they all are of this type, or whether some of them may be actual *šawāhid*. Within chapters, he seems to aim for a roughly even distribution of the three main sources to illustrate his points, a practice that can be seen to strengthen the value of his conclusions across the genres in question and grants his thesis added scientific credibility (Baalbaki 2008: 37-38).

#### 2.5.1.2.2 Scope of his research

When it came to the collected data of the Arabs’ speech (*kalām*, also translated as ‘prose’), its enormity and variety presented a problem of its own, viz. the frequent irreconcilable and idiosyncratic usage that existed among the tribal dialects. Sībawayhi’s approach to the problem is to focus on the most widely attested forms, and in cases where he quotes a certain dialectal *šāhid* unique to a specific tribe, the latter is identified. His main, though not exclusive, sources of dialectal *šawāhid* are known to have been the Ḥijāzī and Tamīmī tribes which are cited 72 and 67 times respectively. The methodological considerations underlying

these choices are obvious, in that the former dialect had belonged to the Prophet and the latter dominated in the southern Iraqi area in which Sībawayhi lived and moved, and constituting the centre of Arabic linguistic studies at the time. Over and above specific cases he does not show a general preference for one dialect over the other (Baalbaki 2008: 37-39). They can in this view be understood to represent an axis, so to speak, with its Western and Eastern poles, around which were clustered the most current or *fuṣḥā* ('purer'/'more correct') dialects of Arabic. Besides systematically narrowing down the scope of relevant material upon which to build his grammar, Sībawayhi demonstrates clearly his pragmatic attitude by his exclusion of certain incompatible *šawāhid* from his analyses, which by their unorthodoxy would have undermined the general rules he was establishing (Baalbaki 2008: 40).

Another significant factor for assessing the value of both prose and poetry as arbitrators of linguistic norms was the period of its production, referred to as '*uṣūr al-iḥtijāj* 'epochs (during which usage could be accepted as linguistic) evidence'. In Sībawayhi's time prose was generally accepted (and only later did its acceptance become limited), yet poetry was subject to greater restrictions. It was generally considered that poetry produced in the so-called *muwallad* period should be rejected as *šawāhid*. This categorization was part of the larger temporal classification of poetry into the four periods of the *jāhilīyūn* (pre-Islamic), *muḥaḍramūn* (straddling both *jāhilīya* and Islam), '*islāmīyūn* (Islamic), and *muwalladūn* (loosely those not of "pure" Arabic descent, also referred to as *muḥdaṭūn* 'moderns'/'innovators'). The latter would include poetry roughly from the second half of the second/eighth century onwards. Sībawayhi likewise largely excludes *muwallad* poetry as *šawāhid* (Baalbaki 2008: 41-42).

### 2.5.1.3 Analytical concepts and methodological approaches

#### 2.5.1.3.1 Methodological considerations related to *šī'r* and *kalām*

A problem related to poetry was the poetic license that poets often made use of which allowed for peculiar constructions that did not otherwise occur in the language. Despite the specific problem this presented for their integration into Sībawayhi's general language system, they found their parallel in the genre of *kalām* as seen in the many incompatible expressions that existed among the tribal dialects. He was thus able to use a similar approach to the categorization and analysis of both genres and to subsume them under the unified theory he

was building. The overall norms of usage were generally extracted from *kalām* against which all source material, including poetry, was measured; an example being Sībawayhi's description of unusual poetic forms in normal speech as being either inadmissible (*lā yajūz*) or weak (*da'īf*). Yet the domination of *ši'r* material as used for *šawāhid* in the *Kitāb* confirms the primacy of this mode of expression as a defining factor for the cultural and linguistic identity of the Arabs. The two genres are complementary; a fact that is clearly reflected in the essentially egalitarian treatment afforded them in Sībawayhi's work. At the same time he is careful to clarify the crucial differences existing between them wherever the distinction is significant to his discussion (Baalbaki 2008: 43, 45-46).

The *šawāhid* available to Sībawayhi was also characterized by a significant presence of *ġarīb* 'unusual'/'rare' material, which reflected the early philologists' special interest in this otherwise marginal field as already mentioned in connection with the early word lists (cf. section 2.2.5). This vast category of literature could likewise hardly be overlooked by Sībawayhi in his analysis; neither did he show any apprehension for the implications of its irregularities to the workability of his theory. On the contrary, he makes the most of the opportunity thus presented him to demonstrate its applicability to even the most abstruse examples of linguistic usage, in the process enriching the diversity of his sources and expanding the basis upon which the theory would be constructed (Baalbaki 2008: 43, 44-45).

The workability of Sībawayhi's system can be adjudged by its effectiveness, as we shall see, for discriminating supposed foreign vocabulary items in the language.

#### **2.5.1.3.2 The concept of analogy (*qiyās*)**

The idea of identifying similar patterns across different instances of linguistic expression, and of extracting the rules governing these specific patterns constitutes the basis for the process of *qiyās* 'analogical extension' as used in Arabic linguistics. Sībawayhi describes it in terms of a speaker's recognition of a certain likeness between two otherwise different items so as to extend analogically to one of them a feature possessed by the other. It can be said to represent the most fundamental concept in the whole of his grammatical theory for its ability to explain the instantiation of forms, patterns, constructions, etc. as well as the presumed inherent logic underlying variegated linguistic phenomena.

An example of the application of this principle can be seen in the ascription to the Ḥijāzis of using the accusative in the predicate of *mā*, as in *mā zaydun munṭaliqan*, on the

basis of their understanding of a similar meaning between *mā* and *laysa*, both of which indicate negation. They have thus most probably afforded the predicate of *mā* the same treatment as the predicate of *laysa*, hence the accusative despite their differences in other areas such as *laysa*, in Sībawayhi's view a verb to which pronouns may be suffixed (e.g. *lasta*, *lastunna* etc.); whereas *mā* is neither a verb nor can pronouns be suffixed to it. The Tamīmīs, on the other hand, use the nominative in the predicate of *mā*, as in *mā zaydun munṭaliqun*, and thus do not effect a change from the supposed original sentence *zaydun munṭaliqun*. Sībawayhi likewise attributes the Tamīmīs' reason for their usage to *qiyās* in that they perceive in the characteristics of *mā* a similarity to other particles like *ammā* and *hal*, and have thus equalized the syntactical behaviour of these particles in relation to their predicate by leaving the case ending unchanged (Baalbaki 2008: 47-48).

This example of the use of the rule of *qiyās* in syntactical analysis finds a parallel in its application to issues of morphology, which was a key indicator for identifying presumed loan words in the language (as described in the next section).

## **2.5.1.4 Characterizing the morphology and phonology of words in Arabic**

### **2.5.1.4.1 Morphology**

Following is a summary of Sībawayhi's theory on morphology and phonology. Regarding morphology, his system can be divided into sections on (1) the root system and (2) form and function (Carter 2004: 99). The following points are relevant to the question of distinguishing foreign elements in the language.

#### **2.5.1.4.1.1 The root system of Arabic words**

The word consists of its root consonants ('radicals') in combination, and the pattern ('structure') in which these consonants are embedded. The pattern may involve one or more vowels together with additional, non-radical consonants or long vowels, namely the ten consonants and semi-vowels *ā*, *h*, *l*, *m*, *n*, *s*, *t*, *w* and *y*. As already mentioned, a convention was used in which word patterns were symbolized generically by the consonants of the verb *fa'ala* (*f*, ' and *l* being its radicals), so that any word whose radicals could be identified, could be transcribed using these same radicals in that combination, radicals being systematically

added or subtracted according to the number of radicals in each particular word. The radicals of the word *qalb* are *q - l - b* which, when replaced by the generic ones *f - ' - l*, show the pattern to be *fa'l*. The same three radicals (in bold) make up the word *mustaq**bal***, whose pattern is *mustaf'al*, of which the remaining parts are augments (*zawā'id*), namely prefixes, infixes and suffixes (Carter 2004: 100-1). This system is now generally taken for granted in Arabic linguistics, and is seen at work in a number of the entries in *al-Mu'arrab*, as in the discussion regarding the supposed pattern for the name of the medieval city Arrajān in southwestern Iran:

*Arrajān*: ...the name of a region, Persian. Abū 'Alī said: Its pattern is *fa'alān* and it is not made into *af'alān*... (Al-Jawālīqī 78)

Likewise in the case of the name 'ayyūb 'Job':

Abū 'Alī said: The analogy (*qiyās*) of the *hamza* of 'ayyūb is that of a radical ('aṣl) and not of an augment (*zā'ida*) since [the word] does not disqualify (*yaḥlū*) to be [of the pattern] *fay'ūl* or *fa'ūl*. So if I rendered it *fay'ūl*, [the word] would be after the analogy - were it Arabic – of its stemming from *al-'awb* as in *qayyūm*. It could also [have been after the pattern] *fa'ūl* as in *saffūd* and *kallūb*, though it was not known in these forms, even as (*li-'annah*) it is not an unusual occurrence that the non-Arab should produce [words] after forms not [found] in Arabic. (Al-Jawālīqī 62)

Further to this system, the following points are relevant to the present discussion:

- (i) The number of roots is finite. This represents the possible combinations of the twenty-eight consonants in roots consisting of one to five consonants, with the exclusion of certain impossible and unused combinations.
- (ii) The number of patterns is also finite, said by one scholar, al-Zubaydī, to be 380.
- (iii) Certain internal restrictions to the pattern system can be identified, which are limited, for instance, by the number of radical consonants (maximum five) and the varying frequency of occurrence of words with different radical counts (three being in

great majority). All quadrilateral and quinquilateral roots also contain at least one liquid or continuant consonant.

(iv) The distribution of radicals and augments is systematic. The consonantal augments generally consist of the same letters that the radicals consist of, and these dual functions are distinguished by their respective ‘places’ *mawāḍi‘* in the word. The three radicals of the word *mamlaka* are *m* - *l* - *k*, where the place of the first *m* indicates its status (*manzila*) as a consonantal augment, and the place of the second *m* its status as radical. Certain generalizations follow. For instance, since the maximum number of consonants in a word is seven, and there can be at most five radicals, one or two of a word containing six or seven consonants must be augments, e.g. *salsabīl* in which the augment is -y- (e.g. *salsabīl*). The same would go for *za‘farān*, where the augment is -ā- (e.g. *za‘farān*), a word al-Jawālīqī mentions as being ‘*arabī ṣaḥīḥ*’ ‘genuine Arabic’ (al-Jawālīqī 221). (Carter 2004: 101-2)

The following entry for the supposed loan word *al-manjanīq* in *al-Mu‘arrab* illustrates a discussion related to the application of the principle outlined in (iv) above:

*Al-manjanīq*, the family of Arabic [speaking tribes] differed on it: The *mīm* is appended. Others said: Rather it is part of the original [root]. Ibn Bundār informed us from Ibn Rizma from Abū Sa‘īd from Ibn Durayd, saying: Abū Ḥātim informed us from Abū ‘Ubayd, saying: I asked a Bedouin regarding *ḥurūb*, whether it belonged to them? He said: Among us it was *ḥurūb ‘ūn*, gouged out eyes, at times *tujnaqu*, at times *nuršaqu*. Thus their utterance *tujnaqu* was meaningful although the *mīm* was appended, and had it been part of the original [root] they would have said *numajnaq*. Al-Māzinī said: The *mīm* is part of the word itself, and the *nūn* is appended, so their utterance is *majānīq*... It is said *manjanīq* and *minjanīq*... It was said that the initial *mīm* and *nūn* are both part of the original [root]. And it was said: They are both appended. And it was said: The *mīm* is part of the original [root] and the *nūn* was appended. It is Arabicized foreign. Al-Farrā’ related *manjūq* with a *wāw*. And another related *manjalīq*. And *qad janaqa al-manjanīq*. And it is said *jannaq*... (Al-Jawālīqī 353-55)



#### 2.5.1.4.1.2 Morphological patterns of words

Certain patterns tend to be associated with certain classes of meaning. Thus the pattern *maf'al* denotes the place of occurrence, *mif'al* the instrument, *fa'āl* the practitioner of a craft and also an exaggerated quality, *'af'al* colours and bodily defects, etc. Certain patterns are also associated with other classes, e.g. the verbal pattern *fa'ula* being intransitive, the pattern *mafā'il* being only found as plurals, etc.

These principles form the basis for classifying the entire etymological and lexical system, nearly all nouns and verbs being accountable to a certain root in a certain pattern. The 'particles' (*hurūf*) do not belong to this system (Carter 2004: 99-104).

#### 2.5.1.4.1.3 Syllable structure and its consequences for morphology

The pattern system implies the principle that syllables have a limited structure. Arabic has only two regular syllable patterns, CV and CVC. Persian, from which most loan words are said to have originated, by comparison has one more, CVCC. The latter pattern may, as will be seen in what follows, occur in Arabic as well, thus disqualifying syllable pattern as a reliable criterion for differentiating loan words of Persian origin, although it would be a criterion for words that have entered from certain other languages.<sup>10</sup> As an aid for analysing such potential differences, and for the purpose of appreciating the underlying structures of certain outwardly irregular morphological forms in Arabic, we will outline some of the rules governing syllable formation in the language as related to nouns (since no verbs are known to have been imported as-is into the language in classical Arabic), excluding the rules related to the definite article *al-* (which is irrelevant to the issue of foreign vocabulary):

The long vowels *ī* and *ū* are analysed as short vowels followed by the consonants *y* and *w* respectively (hence *ī* = *iy* and *ū* = *uw*). All syllables with these long vowels are thus of the CVC type, e.g. *fī* 'in' = *fīy*, *dū* 'possessor of' = *dūw*, the same rule applying to the two diphthongs in Arabic *aw* and *ay*, e.g. *law* 'if' and *kay* 'so that', likewise having the syllable

---

<sup>10</sup> This could be a subject for future studies. In our age, English, for instance, with a potential syllable structure of (C(C(C(((V)(C(C(C(C))))))), could be differentiated on this basis. The minimum syllable length is here 1 (a single vowel), while the maximum is 8 (CCCVCCCC).

structure CVC. By convention the long vowel *ā* is included in this category, the lengthening consonant being seen as a *hamza*, e.g. *mā* ‘what’ = *ma*’, also with the CVC structure.

The general exception to the rule of the CV and CVC patterns, which do not allow either a syllable to begin with a vowel, or to begin or end with a cluster of two or more consonants, is in the case of long vowels, e.g. *ḥārrun* ‘hot’, which corresponds to *ḥa’r* (CVCC) + *run* (CVC), hence CVCC + CVC. This is only allowed with doubled consonants (here *r* + *r*) or in the otherwise common occurrence of the two regular syllable patterns preceding a word already beginning with two consonants. (Carter 2004: 109-10).

#### 2.5.1.4.1.4 Other morphological processes that can potentially affect loan words

The following description of various other morphological processes will be relevant for our purposes:

Pause (*waqf*, lit. ‘stopping’) is the curtailment of a word ending at the termination of an utterance, phrase or breath-group. It most often consists of the elision of final short vowels (as in *muslimūna* > *muslimūn* ‘Muslims’) and the reduction of the feminine suffix *–at* + inflections to *–ah*, so *al-madīn-at-u* ‘the city’ becomes *al-madīn-a(h)*, the *–h* not being pronounced. When the ending is a long vowel it is usually shortened, as when *qāḍī* ‘judge’ becomes *qāḍi* or even *qāḍ*, the latter form being less common.

The pronunciation of words in rhyme, where syllables may be dropped or added for metrical reasons, is related to pause but uses a different mechanism. A word like *šajarun* ‘a tree’ can thus in rhyme be recited as *šajar*, *šajarū* or *šajarun*, whereas only *šajar* is possible in pause.

Substitution (*badal*) is the regular replacement of one sound by another, usually unrelated one. Thus in *qaḍā* ‘judgment’, whose radicals are *q* - *ḍ* - *y*, the final radical *y* is substituted by ’ so as to maintain the pattern *fa’āl* (the sequence *\*qaḍāy*, if inflected, would have produced the unacceptable *\*–āyu-*, *\*–āyi-*, see above).

The glottal stop (*hamza*) can function both as a radical and an augment and is often weakened or elided (Carter 2004: 112-13).

These various operations testify to the existence of rules governing the flexibility of Arabic word forms, further examples of which will be demonstrated below in relation to phonological changes that supposed foreign vocabulary items undergo in the course of being adopted into Arabic.

## 2.5.1.4.2 Phonology

### 2.5.1.4.2.1 Changes occurring in the phonology of loan words

Regarding phonology, Sībawayhi discusses sound changes or *badal* ‘substitutions’ that occur with loan words. He states:

They (the Bedouin) change the foreign words if they do not have them in their own language at all, and sometimes they adapt them to their own patterns and sometimes not. (Carter 2004: 107)

In the introduction of *al-Mu‘arrab*—under the heading *ma‘rifa maḏāhib al-‘arab fī isti‘māl al-a‘jamī* ‘Knowledge of the customs of the Arabs regarding their use of foreignisms’—al-Jawālīqī discusses the background for the Arabs’ usage and adoption of foreign vocabulary in Arabic, summarizing some of the mental and methodological processes and operations involved. He forthrightly acknowledges and validates the existence of loan words in the language, and thus effectively forestalls the need for further speculation about a question that in the past has caused considerable debate on theological grounds:

You should know that there are many who have ventured to modify foreign names when using them. They have exchanged letters that were not from among their own for those closest to them in articulation. Sometimes they exchanged them for letters whose sound was far removed from the originals as well.

[Phonetic] exchange is certainly necessary. Otherwise they would have inserted something not found among their own letters. Sometimes they exchanged the pattern of Persian words for Arabic patterns. This latter modification was effected by substitution letter for letter, by adding or omitting a letter, by modifying the pointing of the letters, by silencing letter-pointing [with *sukūt*], or pointing such silence. Sometimes they would leave the letter as it was without modifying it.

Among the letters they modified were those falling between *jīm* and *kāf*, which they sometimes rendered with *jim*, sometimes with *kāf*, and sometimes with *qāf*, due to the proximity of *qāf* to *kāf*. Some would say *kurbaj*, and others *qurbaq*.

Abū ‘Amrūwīn once said: “I heard al-Aṣma‘īya state: it is the passage that has been stated to him: *kurbak*. He said: They prefer *kurbaj*. Sālim ibn Quḥfān said regarding *qurbaq*:

I did not drink after the crow’s (*qurbaq*) ṭawī...

Likewise they say: *kīlaja*, *kīlaqa* and *qīlaqa*; *jurbuz* for *al-kurbuz*; *jawrab*, its origin being *kawrab*; *mūzaj*, its original being *mūzah*. They also modified the letter falling between *bā’* and *fā’* to *fā’*, and sometimes to *bā’*. They said *fālūd* and *firind*, and some said *birind*.

They exchanged *sīn* for *šīn*, and said, in reference to the desert (*ṣaḥrā’*), *dast*, while in Persian it is *dašt*. They said *sarāwīl* and *’ismā’īl*, their originals being *šarwāl* and *’išmāwīl*, due to the proximity of *sīn* to *šīn* when muttering (*fī al-hams*).

They substituted the *lām* for the *zāy* in *qafšalīl* which signifies ‘ladle’ (*mağrafa*), its original being *kafjalāz*, and rendered its *kāf* as *qāf*. Its *jīm* they rendered *šīn*, the *fathā* a *kasra*, and the *alif* a *yā’*. Among others they changed the vowels in *zūr* and *āšūb*.

Among [the words] whose [respective] constructions (*’abniyatihim*) were appended to [or ‘used as a basis for various Arabicized forms’] (*’alḥaqūhu*) were: To *dirham* they appended *–hajra’in*, and to *bahraj* *–salhab*. To *dīnār* they appended *–dīmās*, and to *’ishāq* *–’ibhām*. To *ya’qūb* was appended *–yarbū’*, and to *jawrab* *–kawkab*. To *šubāriq* *–’aḍāfir*, and to *ruzdāq* *–qurṭās*.

Among the foreign words that were extended or shortened we find *’ibraysam*, *’isrāfīl*, *fayrūz*, and *qahramān*, the original of the latter being *qirmān*.

Among the words that were left in their original state, unchanged, are *ḥurasān*, *ḥurram*, and *kurkam*. (Al-Jawālīqī 54-56)

#### 2.5.1.4.2.2 Basic assumptions regarding phonology

Sībawayhi describes the processes of assimilation of loan words in terms of a set of postulates and general principles of which the following are particularly relevant to our subject:

## Codifying the collected corpus of the Arabic language

- (a) There is a general principle of least effort. For example, there is a reluctance to repeat the same sounds too closely together, which is usually avoided by combining the two separate sounds into one of double length, e.g. *radada* 'to return' becomes *radda*.
- (b) The set of sounds peculiar to one language (phonemes) is finite: in Arabic it is twenty-nine (counting *ā* and ' as two), which may have allophones.
- (c) Certain combinations of sounds have not been observed, e.g. consonantal sequences such as [nr], [nl], vocalic sequences [iu], [ui] etc. (cf. section 4.1.1.1.3 below).
- (d) Sound changes can be constrained by the need to avoid creating homonyms, thus *watada* does not become \**wadada* > *wadda* because there is already a word *wadda* even though the sound change [td] > [dd] itself is attested in other words.
- (e) Words borrowed from foreign languages will be adapted to Arabic phonology (and morphology, see above) (Carter 2004: 122).

### 2.5.1.4.2.2.1 Disallowed or disapproved sound combinations in Arabic

The following are examples of entries in *al-Mu'arrab* where principle (c) above is seen at work:

In the *ḥadīṭ* of al-Qāsim ibn Mujaḥmira he said: ....*al-iṣṭaflīna*....not in pure Arabic, since *ṣād* and *ṭā'* are hardly ever (*lā yakād*) combined [in a word], though (*wa-innamā*) it occurs in *al-ṣirāṭ* and *al-uṣṭumm* since their originals are [with] *sīn*... Ibn al-A'rābī said: "al-iṣṭaflīn" ...*ša'āmīya* dialect... it is also water. (p. 92)

From the term *al-bustān* is this called *bast*, and no reliable sources ever pronounced a word of the Arabs consisting of *bā'*, *sīn* and *tā'*. (p. 102)

*Bağdād*: A foreign name. As though *bağ* is an idol/image. *Dād* is a gift. So it is as though a gift of an idol.

Al-Aṣmā'ī had disliked saying *bağdād* with *dāl* and *ḍāl*, (but rather) *bağdān* with *nūn*. And *mağdān* with *mīm* in place of *bā*'.

The Arabs have given word to it...

Abū Ḥatīm said: I asked al-Aṣmā'ī regarding *bağdād*, *bağdāḍ*, *bağdān* and *bağdīn*: Is all this given word to? He disliked having to say anything about it, and said: this is ugly (*radī'*), I fear that it should be idolatry (*širk*), and he said: the most hateful to me is with *ḍāl*..., and it was said the city of peace (*madīnat al-salām*)... (p. 121-2)

Abū Bakr said: *al-zanr*. Obsolete verb... I do not reckon it to be Arabic... Sībawayhi said: There is not in the speech of the Arabs an unvowelled *nūn* followed by *rā*', such as *qanr* or *zanr*. (p. 220)

*Al-šārūj*: lime (*nūra*) and its ingredients/mixture that *tušarraju bi-hā* the pools and the pigeons. It is said *šarrajtu* the pool: when you covered it with clay. And *al-šārūj*: Arabicized Persian. And likewise every word containing *šād* and *jīm*, since the two do not coincide within one word in the Arabs' speech. (p. 261)

*Al-qubj*: partridge. Arabicized Persian. Since the *qāf* and the *jīm* are not combined in the same word in Arabic speech... (p. 309)

He said: *al-qāquzza*: One of the containers for drinks/wine (*al-šarāb*). It is *al-qāqūza* and *al-qāzūza* also. It is said that it is Arabicized. There is not in the Arabs' speech what divides [with an] *alif* between the two same letters that stem from the construction *qafz* and its like. (p. 321-2)

Likewise *al-narjis*: Arabicized foreign, and the grammarians have mentioned it in constructions, and it does not correspond to anything in speech. If a construction like *fi 'lil* occurred in the poetry of old, reject it, it is contrived. And if the half-breed (*muwallad* indigenous non-Arabs) made this construction and used it in a poem or utterance, then the former rejected it. There does not occur a name in the Arabs' speech with a *nūn* followed by *rā*'. (p. 379-80)

Al-Layth said in the utterance of Ru'ba: "He prepared *aḥṭāl* for him and *narmaq*."

*Al-narmaq* is Arabicized Persian... Another said: Its meaning is *narm* and it is 'good'...

And it is related from him: *al-narammaq* signified a soft white garment, which is *narmah* in Persian, resembling a mirage (*sarāb*)... (p. 381-2)

Al-Azharī related from Ibn Durayd: *al-narja*: Piece of wood that overturns the earth. And in the rare inflections: *al-nawraj*: mirage (*al-sarāb*). And *al-nawraj*: path (*sikka*) of the plowman (*ḥarrāt*). Al-Layth said: *al-nayraj*... All of this is imported, since the *nūn* and *rā'* are not combined in one [and the same] word in the Arabs' speech. (p. 385)

#### 2.5.1.4.2.2 Tracing sound changes in loan words

Following are examples of entries in *al-Mu'arrab* that indicate changes in phonology undergone by a word of foreign origin in relation to a supposed original in the foreign language:

*Al-istār*: Abū Sa'īd said: I heard the Arabs say *istār* for four since it is *jihār* in Persian, so they Arabicized it saying *istār*.

Jarīr said: ...

Al-A'shā said...

This is the pattern that is attributed to *istār*... (p. 90-1)

Ibn Durayd and Ibn Qutayba said: *al-bahraj*: *al-bāṭal*. In Persian it is *nahrah*. Al-'Ajjāj recited...

Ibn Durayd said...

Al-Azharī said: *al-bahraj* (from the verse recited above by Ibn Durayd) is not pure Arabic, its original being *nabharaj* which is *al-radī'* of dirhams, as though its original were *nuwwāra*, so it was said *nabahraj* and *bahraj*. ... (p. 96-7)

*Al-bartūlla*: A Nabatean word, not from the Arabs' speech.

Abū Ḥātim said: Al-Aṣmā'ī said: *bar* ibn. The Nabateans rendered the *zā'* as *ṭā'*.

As was thought, they intended *Ibn aḍ-Ḍill*, but did they not see that they say *al-nāṭūr* while it is actually *al-nāẓūr*. (p. 116)

It is said regarding *al-tūt*: It is Arabicized Persian. Its original is *al-tūt* which the Arabs Arabicized, rendering the *ṭā'* as *tā'*, and appended it to some of their constructions. (p. 138)

As to *al-ḥubb* into which is deposited water, it is Arabicized Persian, it is not truly old Arabic (*muwallad*). Abū Ḥātim said: Its origin is *ḥunb* which was Arabicized, thus they substituted the *ḥā'* and dropped the *nūn*, so they said *ḥubb*. And a man is called *ḥunbī* because they withdrew to their dear ones (*al-aḥbāb*)... (p. 168)

*Al-dast*: the desert. In Persian it is *dašt*... (p. 186)

In the *ḥadīt* as well as in [what was] sent [down] (*al-mab'at*): The angel of the knife *darahraha*. Ibn al-A'rābī said: It [has] a bowed (*mu'wajj*) head, which the generality call *al-minjal*. Its original is from Persian *darah* which the Arabs Arabicized and added to it letters of their kind, and so they did, as they said *muqamjar* of the bowmaker, and *baraq* and *baḍaj* for a load. (p. 199)

Ja'far ibn Aḥmad informed us from 'Abd al-Bāqī ibn Fāris from ibn Ḥasnūn from Ibn 'Uzayr in His exalted utterance [Qur'ān]: *ṭūbā lahum*. He said: It is said *ṭūbā*: the name for a garden in Hindi (*al-hindīya*). And it is said *ṭūbā*: a tree in the garden. The grammarians call it *fu'lā* from goodness (*al-ṭīb*). This is the utterance. The original of *ṭūbā* is *ṭūybā*, so they changed the *yā'* to a *ḍamma*, preceded by *wāw*. (p. 274)

*Qābūs*: A foreign name. In Persian it is *kāvūs*, which was Arabicized so it is said *qābūs* so as to agree with Arabic. Al-Nu'mān ibn al-Mandir was designated *Abū Qābūs*...

The absence of its inflection is evidence that it is foreign, for had it been [derived] from the term *al-qabas* it would have been inflected, even as had you named a man –



*‘āqūl* you would have inflected [it]. Ḥujrun Khālīd said: “...as the activity of *abī qābūs*...”

They stood in need [of it] in the poem, so they reduced it with an apocopation. ‘Amruw ibn Ḥassān said: “...did you see *abā qubays*...?” (p. 307-8)

#### 2.5.1.4.2.3 Method of describing Persian phonemes

Sībawayhi describes the sounds of Persian words in terms of Arabic letters, the Persian [g], for instance, being the “sound between [k] and [j],” which in Arabic is replaced by [j] (or occasionally [q]), e.g. *al-lajām* ‘bridle’ for Persian *ligām*, or the [p] being the “sound between [b] and [f],” which is replaced by [f] or [b] in Arabic, e.g. *parand* becoming *firind* or *birind* ‘kind of cloth or garment.’ (Carter 2004: 107)

#### 2.5.2 Speech as an expression of behaviour

The basis for Sībawayhi’s value system in grammar was that speech is to be judged as a form of behaviour, and he thus judges speech by behavioural criteria. As such, he uses ethical terminology in categorizing the viability of speech acts (Carter 1973: 146-7). Carter concludes regarding the criteria Sībawayhi uses to distinguish this viability:

It is the listener who determines rightness: much of what we say, as Sībawayhi points out, is conditioned by what we think our listener expects, whose questions we continually anticipate. ... "right" utterances are those in which the speaker fulfills his social obligation to communicate... (Carter 1973: 146-7)

Speech is here seen as a function of the speaker’s ability to communicate the desired message or content, and although such a function is inextricably bound up with conventional forms of expression in the language, theoretically it should not exclude the possibility of this function being filled by foreign or unconventional expressions. This train of thought was not, however, Sībawayhi’s main entry point for distinguishing loan words, this being reserved to the morphological and phonological criteria already mentioned.

## Factors affecting the process of canonization of the Arabic language

*Religious, social and political aspects bearing on the endeavors to codify the language.*

---

### 3.1 The role of the Qur'ān and Islamic institutions in the canonization of the language

With the coming and expansion of Islam, Arabic took on the status of a religious, administrative and educational language of an empire consisting of a diversified population. The Qur'ān was at the centre of this expansion and influenced policy both directly and indirectly; the latter through the exegetical (*tafsīr*) activity of Muslim scholars who sought to apply its provisions to different areas of the life of society and the individual.

The holy book asserted in clear terms its unique and inimitability status:

*qul la-'ini ijtama'ati l-'insu wa-l-jinnu 'alā 'an ya'tū bi-miṭli hāḍa l-qur'āni lā ya'tūna bi-miṭlihi wa-law kāna ba'ḍuhum li-ba'ḍin ṣāḥibān (Q 17:88)*

Say: 'If men and jinn banded together to produce the like of this Koran, they would never produce its like, not though they backed one another.' (Arberry)

This notion of a special station allotted to it was understood by extension to encompass the language itself, whose identity was established in unmistakable terms against other languages as Arabic, an '*arabīyun mubīn* 'a clear/manifest Arabic' at that:

*...al-kitāb al-mubīn \* 'innā 'anzalnāhu qur'ānan 'arabīyan... (Q 12:1-2)*

...the Manifest Book. We have sent it down as an Arabic Koran...

...*'anzalnāhu ḥukman 'arabīyan...* (Q 13:37)

...We have sent it down as an Arabic judgment.

*wa-mā 'arsalnā min rasūlin 'illā bi-lisāni qawmihi li-yubayyina lahum...* (Q 14:4)

And We have sent no Messenger save with the tongue of his people, that he might make all clear to them...

*wa-laqad na'lamū 'annahum yaqūlūna 'innamā yu'allimuhu bašarun lisānu llaḏḥī*

*yulhidūna 'ilayhi 'a'jamīyun wa-hāḏā lisānun 'arabīyun mubīn* (Q 16:103)

And We know very well that they say, 'Only a mortal is teaching him.' The speech of him at whom they hint is barbarous; and this is speech Arabic, manifest.

*bi-lisānin 'arabīyin mubīn... wa-law nazzalnāhu 'alā ba'ḏi l-'a'jamīna \*fa-qara'ahu 'alayhim mā kanū bi-hi mu'minīna* (Q 26:195, 198-99)

...in a clear, Arabic tongue... If We had sent it down on a barbarian and he had recited it to them, they would not have believed in it.

*qur'ānan 'arabīyan ḡayra ḏī 'iwajin la'allahum yattaqūna* (Q 39:28)

An Arabic Koran, free from tortuous wording, to the intent that they may fear God.

*kitābun fuṣṣilat 'ayātuhu qur'ānan 'arabīyan li-qawmin ya'lamūna* (Q 41:3)

A Book whose signs have been distinguished as an Arabic Koran for a people having knowledge...

*wa-law ja'alnāhu qur'ānan 'a'jamīyan la-qalū lawlā fuṣṣilat 'ayātuhu 'a'a'jamīyun wa-'arabīyun...* (Q 41:44)

If We had made it a barbarous Koran, they would have said, 'Why are its signs not distinguished? What, barbarous and Arabic?'

*wa-min qablihi kitābu mūsā 'imāman wa-raḥmatan wa-hāḏā kitābun muṣaddiqun*

*lisānan 'arabīyan li-yundīra llaḏīna ḡalamū wa-buṣrā li-l-muḥsinīna* (Q 46:12)

Yet before it was the Book of Moses for a model and a mercy; and this is a Book confirming, in Arabic tongue, to warn the evildoers, and good tidings to the good-doers.<sup>11</sup>

The notion of foreignness is unmistakable in verses 16:103 and 41:44, upon which Carter comments:

[T]here is an implicit contrast between “Arabic” *‘arabī*, and “non-Arabic, foreign” *‘a‘jamī*. Interpretations differ on the circumstances and meaning of these...two verses, but the implications are unmistakable: this revelation was delivered in a language familiar to its audience. (Carter 2006: 120)

Much has been said regarding the possibility of the existence of foreign vocabulary in the Qur’ān, but the question has remained unresolved. Al-Suyūṭī provides a summary in his *Itqān* of the discussion (*‘imām* would here signify ‘linguistic authority’):

The Imāms differ as to the occurrence of foreign words in the Qur’ān, but the majority, whom are the Imām al-Shāfi‘ī, Ibn Jarīr, Abū ‘Ubayda, the Qāḍī Abū Bakr, and Ibn Fāris, are against their occurrence. (Jeffery 1938: 6)

The conception of the inimitability and Arabness of the language of the Qur’ān combined to imbue the Arabic language itself with a halo of sacredness, an aura which particularly the Arab Muslims sought to extend to many aspects of their own tribal and literary cultural heritage. The result was the diffusion of certain unwarranted assumptions about the insular and exceptional character of the language, although this could rightly be said to apply to aspects of it such as the regularity and flexibility of the derivational system in Arabic grammar.

---

<sup>11</sup> Translations by Arberry, except 39:28 by Rodwell.

### 3.2 Social factors influencing the canonization of the language

The Arab conquests have been described as the engine leading to linguistic mixing and levelling of the various dialects of the Arab tribes. The resulting waves of migration into the conquered territories and the interaction that took place with heretofore unacquainted peoples—each with an unknown language and dialect—called for the establishment of a common written, if not always spoken medium of communication. The original Arabic spoken by the Arabs, and in particular the Bedouin Arabs, came under pressure for change from the influence of these foreign languages. At the same time a conscious effort was being made on the part of the Muslims to codify and preserve the Arabic language as handed down in the Qur’ān, in the pre-Islamic poetry, and as spoken by the Bedouins, whose language had not been influenced to the same degree as those in the forefront of the Muslim advance (Suleiman EALL Vol. I: 174). The expression used most frequently among the Muslim grammarians to refer to the Arabic language is, incidentally, *kalām al-‘arab*, the language of the Bedouin (Versteegh 1996: 15). Local linguistic varieties of the language, described by the grammarians as *lahn* ‘solecism’, were perceived as a threat to what was regarded as the “pristine purity” of Arabic as represented by the recognized sources (cf. sections 2.5.1.2-2.5.1.2.2 above) and termed the *fuṣṣḥā* variety of the language, its standard form (Suleiman 174).

There are several references to Bedouins in *al-Mu‘arrab*, referred to as either *‘a‘rābī* or *bādiya* in their designated role as guardians of original Arabic. It demonstrates again the importance, respect and value accorded their cultural capital and linguistic perception, as a result of which they are called upon as guides for correct usage. The following are examples:

Al-Layth: *al-dāshan*: Arabicized, and not from the speech of the Bedouins (*al-bādiya*)... (p. 193)

*Al-zanfālīja*, and it is said *al-zanfīlija*: Arabicized foreign. Al-Aṣma‘ī said: I heard it from the Bedouins. Abū Ḥātim said: I heard it from the mother of al-Haysham and others besides her, fluent in their speech, as though they transformed it into their (own) speech. Al-Aṣma‘ī said: It is Persian *zīn fālah*: container. (p. 218)

Al-Aṣma'ī said: It is said tamr *sihrīz* and *šihrīz*. He said: I heard a Bedouin ('a'rābī) say *šuhrīz*... It is Arabicized Persian... (p. 247)

*Al-qurm*: a type of tree. Abū Bakr said: I do not know whether it is Bedouin ('a'rābī) or an import. (p. 317)

*Al-manjanīq*, the family of Arabic [tribes] differed on it: The *mīm* is appended. Others said: Rather it is part of the original [root]... Abū Ḥātim informed us from Abū 'Ubayd, saying: I asked a Bedouin regarding *ḥurūb*, whether it belonged to them? He said: Among us it was *ḥurūb* 'ūn, 'gougged out eyes', at times *tujnaqu*, at times *nuršaqu*. Thus their utterance *tujnaqu* was meaningful although the *mīm* was appended, and had it been part of the original [root] they would have said *numajnaq*... (p. 353)

A political factor that created some controversy in relation to the question of foreign vocabulary in Arabic was the *šu'ūbīya* movement, which found its beginnings among the educated classes of Persians—though including other nationalities as well—in the second/eighth century, its activities peaking in the third/ninth century. The Arabs' exaggerated veneration for their own language at the expense of the converts' in the Persian speaking territories provoked a reaction among these peoples in favor of their own tongue. The proponents of the *šu'ūbīya* movement found it necessary to vindicate the excellences of the long-standing Persian literary heritage of which they were the bearers, and did so by criticizing elements of the Arab-Islamic tradition and by attributing to foreign origin as many Arabic words as possible. The counter movement—the anti-*šu'ūbīya*—on the other hand, tried to weaken the arguments presented for labeling Arabic words as loans. A famous anti-*šu'ūbī*, despite his Persian background, was the lexicographer Ibn Fāris, who in his influential dictionary *al-Muǧmal* frequently states *wa-huwa 'arabī maḥḍ* 'it is pure Arabic' (Enderwitz EI Vol. XI: 513-14; Kopf 1961: 201-2). The reverse expression *laysa bi-'arabī maḥḍ* 'not in pure Arabic,' *lā 'aḥsibuhu 'arabīyan maḥḍan* 'I do not reckon it to be pure Arabic,' etc. occurs a score or so times in *al-Mu'arrab* as exemplified by the three entries for *al-bāsana* (a presumed place), *al-šiṣṣ* 'sharp thief' and *al-ṭaraš* 'deafness' (Hava s.v.) below:

In the *ḥadīṭ*: Adam came down from Paradise at *al-bāsana*. It is said that it is a forged verse. It is not pure Arabic. (Al-Jawālīqī 131)

As to *al-šiṣṣ*, Ibn Durayd said: I do not reckon it to be pure Arabic. (p. 257)

He [Ru'ba] said: As to *al-ṭaraš*, it is not pure Arabic. Rather it is among the words that are not truly old Arabic (*muwallad*). It is a rank of deafness (*ṣamam*) with them. And al-Ḥarbī said: ...I consider it Persian. (p. 272)

The *šu'ūbīya* movement, although significant in its time, had a limited life span and its impact on conceptions about the actual indigenous or foreign identity of Arabic words proved in the end to be relatively inconsequential. It is probably not a coincidence that no original *šu'ūbī* tract has survived, making access to information about the movement's activities accessible only through reconstruction by way of anti-*šu'ūbī* texts that answer to their allegations (Enderwitz 515).

## **The role of identifying foreign vocabulary in the process of canonizing the language**

*The grammarians' viewpoints on the issue of loan words in Arabic and their  
conclusions on the subject. An introduction to and case study of al-Jawālīqī's  
al-Mu'arrab.*

---

### **4.1 Perspectives on and criteria for distinguishing foreign vocabulary in Arabic**

#### **4.1.1 Attitudes of the learned to the question of the existence of foreign vocabulary in the Qur'ān and in the Arabic language in general**

The issue of foreign vocabulary in Arabic first came up in the context of Qur'ānic statements about the Arabic nature of its language (Rippin EQ Vol. III: 226), a notion that came to be colored by the cultural outlook of the Arabs: as the ruling aristocracy of an expanding Islamic empire they harbored sentiments of ethnic superiority in relation to non-Arabs or even to Arabs of lower social-economic status such as those of Iraq and Bahrain (Lewis 2002: 72-3), an attitude we today could label as nationalism (in the meaning of, “a sense of national consciousness exalting one nation above all others and placing primary emphasis on promotion of its culture and interests as opposed to those of other nations” [Merriam-Webster 2009]). It received renewed attention as a result of systematizations in Arabic morphology and phonology that revealed the language's internal structural consistencies as expressed in the works of al-Khalīl and Sībawayhi. The question was also approached from the perspective of rhyming and semantic value, where foreign vocabulary was seen to complement and enrich the expressive potentiality of the language. The pragmatic attitude to the issue that resulted



from the latter consideration is summarized succinctly by al-Jawālīqī with reference to *al-qāfiya* ‘rhyme’ and *mustaṭraf* ‘considered as being unusual’ in the introduction to his work:

Abū Ḥātim recounted: Ra’uba ibn al-‘Ajjāj and the pure-spoken ones (*al-fuṣṣḥā*), such as al-A‘shā and others - : they seized on one of the foreign words for rhyme so as to make [their verse] sound original, but they did not make use of [what would be] considered unusual, nor inflect it, derive any verbs from it, or discard the unusual original. (Al-Jawālīqī 57-58)

With recognition of the advantage of using foreign vocabulary under certain conditions—notably in poetry, and later in the imported sciences—and with increased experience in using it in the language, the comprehension of the dynamics of the naturalization-process for such vocabulary and of the rules governing it which is evident in al-Jawālīqī’s work can be said to have emerged. The acceptance of the idea of its potential benefits elicited, besides, a questioning of the notion of the inherent superiority of the Arabic language and, consequently, of Arab culture in general (in this context the two may be regarded, in effect, as two sides of the same coin). The latter development found extension in the sophisticated philosophical-linguistic speculations originated by the renowned al-Fārābī (d. 339/950) who became known as the “second teacher,” the first being Aristotle. He considered all languages to be based on logic, stating:

[T]his art [of logic] is analogous to the art of grammar, in that the relation of the art of logic to the intellect and the intelligibles is like the relation of the art of grammar to language and expressions. That is, to every rule for expressions which the science of grammar provides us, there is a corresponding [rule] for intelligibles which the science of logic provides us. (Nasr, Leaman 1996: 180)

Based on this conception, he suggested the possibility that foreign languages could be superior to Arabic in certain respects; a case in point being the latter’s inability to express the copula which could be found in all other known languages (Versteegh 1997: 83).

#### 4.1.1.1 Muslim grammarians' reckoning of foreign vocabulary in Arabic

Kopf (1961) has surveyed the criteria used by Muslim philologists for distinguishing loan words in Arabic, stating:

Lacking sufficient knowledge of the respective foreign languages, the Arab (*sic*) philologists resorted to deductive criteria which they determined on the ground of Arabic itself. (Kopf 1961: 197)

He goes on to list these criteria under three headings:

1. *the morphological structure*
2. *the barren roots*
3. *phonetic features*

Regarding these rules it may be noted that loan words sometimes conformed to them as well, so they could not be used indiscriminately.

Criterion 1, of morphological structure, was a result of Sībawayhi's schematizing the morphological patterns of words found in Arabic, making it an easy task to distinguish words that did not conform to these patterns; such were usually marked as being of foreign origin. Barren roots (criterion 2) refers to the existence of only one derivative of a root which would mark the word as a potentially foreign. Even the possible permutations of its radicals was considered a factor in this regard. With regard to phonetic features (criterion 3), al-Khalīl had examined the characteristics of sounds in Arabic and used them as a basis for distinguishing foreign vocabulary. Arabic words having quadri- or quinquiliteral roots contain at least one of the letters *r, l, n, f, b, m*, which he termed *ḥurūf al-ḍalaq wa-l-ṣafawīya* 'smooth and labial [consonantal] letters.' Words found to be at variance with this rule were considered non-Arabic. Another rule that was established was that no Arabic word begins with the two letters *n* and *r*. Furthermore, certain pairs of letters were found not to co-occur in Arabic roots, such as *j - ṣ, t - ṭ, j - q, j - ṭ, j - k; z* after *d* at the end of a word, as in *muhandiz*; and *j - t* in a word containing none of the letters *r, l, n* (*ḍawlaqīya* – from *ḍawlaq* 'tip of the tongue'), as in *jabt*.

Al-Khalīl also conceded that exceptions could be found to the phonetic rules regarding quadri- or quinquiliteral roots; al-Azharī is said to have found two exceptions to one of these rules, and Ibn Durayd five or six to another rule (Kopf 1961: 198-200).

In the chapter of *al-Mu‘arrab* entitled *mā yu‘raf min al-mu‘arrab bi-‘tilāf al-ḥurūf* ‘What is known about Arabicized words in concordance with the letters’ (p. 59), al-Jawālīqī enumerates the phonological rules pertaining to the subject of indentifying foreign vocabulary:

*Jīm* and *qāf* are not combined in Arabic words. So whenever they come together in a word, know that it has been Arabicized. Among them is *jalawyaq*, *jarandaq*, *al-jawq*, *al-qabj* and a man who is *‘ajwaq*. You will see this [phenomenon] explicated in its proper place in what follows.

Nor are *ṣād* and *jīm* combined in Arabic words. For instance in *al-jiṣṣ*, *al-ṣanja*, *ṣawbalaḥān* and the like.

In original Arabic word constructions there are no nouns in which *nūn* is followed by *rā*. So whenever you encounter this you will know that it is an Arabicized name, such as *narjis*, *nars*, *nawraj*, *nirsian* and *narja*, as you will see explicated [in its proper place].

In their speech is not found *zāy* after *dāl* unless it is an import. For instance: *al-handāz* and *al-muhandiz*, in which *zāy* has been exchanged for *sīn*, such that they say *al-muhandis*.

None of the reliable authorities of the Arabic language pronounced any construction made up of *bā’*, *sīn*, and *tā’*. So if this occurs in a word, it is imported.

As to [exemplary] models of Arabic (*‘amṭila al-‘arab*), the most ideal are those consisting of letters whose place of articulation is at a distance from each other.

The weakest of the letters are the smooth consonants (*ḥurūf al-ḍalāqa*), of which there are six. Three are [articulated] at the tip of the tongue, which are *rā’*, *nūn* and *lām*. And three from the two lips, which are *fā’*, *bā’* and *mīm*.

The quadri- and quintiliteral [roots] are thus not free of them, excepting [the case of] *‘asjad*, where it can be said that *sīn* is akin to *nūn* because of its high, thin tone, and the nasal sound of *nūn*.

So if you come by examples of quinti- or quadriliteral [roots] without one or two of the smooth consonants, know that it is not a part of their speech, such as *‘aqjaš*, *ḥuḏāṭaj* and the like.

The [present] synopsis of statements pertaining to this field shall suffice [us].  
(Al-Jawālīqī 59-60)

#### **4.1.1.2 Theological hegemony and the eventual penetration of secular ideas into the Islamic domain**

Attitudes to the possibility of the existence of foreign vocabulary in the Qur’ān among Muslim scholars range from absolute denial to complete acceptance. Denial is usually based on theological considerations as seen in the attitude of Abū ‘Ubayda, who categorically denied the existence of foreign words in the Holy Book (Rippin EQ Vol. III: 229).

The question of foreign vocabulary has always been pressed by theological considerations in Islam. Any pragmatic and scientific contributions in this regard were relegated to a subordinate and contingent status in relation to the ruling religious dogma, depending on the latter for their approval, currency and, ultimately, validity. In the nineteenth century, secular notions began to infiltrate the educated classes in Islamic societies through their increasing contact with educational institutions of the West, and this opened up for a challenge to the long-held theological notions about the sanctified status of the Arabic language in Islam. The status of foreign vocabulary in Arabic was one of the areas that came to be questioned and investigated anew as a result of this intellectual shift, as exemplified by Arthur Jeffery’s (1938) comprehensive research on the foreign vocabulary of the Qur’ān in his book of the same name. Even so, such later developments were understandably almost completely dependent on the work done by their predecessors in Arabic linguistics, especially in the case of lexicography by virtue of the nature of the subject. In the field of grammar, on the other hand, modern linguists such as Ewald, Nöldeke, Brockelmann, and Reckendorf attempted to detach their theories in syntax from those of their forerunners (Talmon 1998: 74).

## 4.2 Al-Jawālīqī, author of *al-Mu‘arrab*

Abū Maṣṣūr Mawḥūb b. Aḥmad b. Muḥammad b. al-Khaḍīr al-Jawālīqī, known as al-Jawālīqī, was born in Baghdad in 1073, and died there in 1144. Brockelmann traces his lineage to an ancient family, which would seem to belie the humble signification of his title al-Jawālīqī, “maker, seller of sacks”, whose Arabicized plural form (sg. *juwālīq*) is derived from the Persian *gowāl(e)* “sack.” His best known work is *al-Mu‘arrab min al-kalām al-‘a‘jamī* which, taken together with the works of his teacher—the celebrated philologist al-Tibrīzī (d. 502/1109) who wrote standard textbooks on *adab* (“belles-lettres,” covering poetry, artistic prose etc.)—is said by Fleisch (EI Vol. II: 490) to have raised the cultural level of the Arabic language from the low state to which it had fallen under Saljuqid rule. It summarizes the works of his predecessors and is recognized to this day as the primary reference work on the subject. He held the chair of philology at the Niẓāmīya of Baghdad, becoming the second successor of al-Tibrīzī to do so. Another of his well known works, *al-Takmila fī mā yalham fī-hi l-‘amma*—the second of the two works of which manuscripts exist—is on the subject of incorrect expressions (Fleisch EI Vol. II: 490). Of him, Ibn Khallikān writes: “[He was] a learned scholar, and master of all branches of literature...A number of instructive works were composed by him and got into wide circulation...” (Haywood 98).

### 4.2.2 *Al-Mu‘arrab*

Al-Jawālīqī’s desire in compiling his work was to preserve for posterity what was known about loan words in Arabic by his time (al-Jawālīqī 51). The few anecdotes about him that have reached us—as much as the contents of his two surviving works—allow us to understand that he was generally concerned about the direction that the contemporary Arabic language was taking, and about its possible fate: the stream of foreign terms that were constantly being introduced into it looked to be gradually undermining the language’s original and unique characteristics as exemplified in the Qur’ān, the pre-Islamic poetry and attested Bedouin speech, and to be threatening its integrity (Fleisch 490; al-Jawālīqī 51).

*Al-Mu‘arrab* helped close one of the few remaining gaps to the complete canonization of the Arabic language in that it collected and synthesized what was known up until his time

about the “impurities” that were considered to have entered the language from abroad. It consists mainly of material collected from other works and the work itself can thus be seen to have capped the aggregate of developments that had taken place on the subject of foreign vocabulary by al-Jawālīqī’s time.

### 4.2.2.1 Contents

#### 4.2.2.1.1 Form and approach

Al-Jawālīqī seems to have had no specific ambitions regarding the *form* of his work, letting it instead conform to the general mode of discourse that characterized learning of the period; it resembles answers given—in an unassuming and matter-of-factly style—to open questions. The diligence with which his points are supported with citations illustrating actual usage wherever possible is reassuring to the reader and adds credibility to the work. Shākir (al-Jawālīqī 1969) has checked the correctness of these many quotations against the available sources, supplying critical comments in footnotes.

#### 4.2.2.1.2 Subject matter

Al-Jawālīqī’s introduction to the book succinctly describes its subject matter, as well as the purpose of and benefits accruing from such a compilation:

This is a book in which we [shall] mention what the Arabs expressed with regard to foreign speech, and what the Qur’ān expressed on it and what was introduced in the reports from the Messenger, his companions and followers. Also what the Arabs mentioned in their poetry and communications. This in order that it may be distinguished between what has been introduced and what is pure [Arabic].

There is significant benefit in this knowledge, which guards the derivative [feature of Arabic grammar] in order that nothing of the Arabic language be [erroneously] attributed to a foreign language.

Abū Bakr ibn as-Sarrāj said in his book on derivation, in the chapter on what the researcher of etymology must guard against and watch out for: “Among the things of which every cautious one should beware [of doing] is to make derivatives in Arabic

of anything [that is actually] from a foreign language, for he would then be in the position of one who claims the bird to be born of the fish.”

It is related on account of Abū ‘Alī: I saw Abū Bakr pondering the expression “būṣī” in order to find its derivatives, so I said: What are you about? Why, it is Persian, nothing but “buzīd,” which is our name for good luck! He said: its meaning is secure. Then Abū Bakr said: You have relieved me. (p. 51-52)

#### 4.2.2.1.3 Standpoints on the issue of the occurrence of loan words in the Qur’ān

He goes on to discuss the differing views that have been expressed on the delicate issue of the existence of foreign vocabulary in the Qur’ān over time, and attempts to lay these long-standing controversies to rest:

As to what has been introduced [of foreign words] into the Qur’ān, the learned ones have differed: Some have said: There is nothing in the book of God besides Arabic.

No less an authority than al-Ḥasan ibn Aḥmad, from Da‘laj from ‘Alī ibn ‘Abdu l-‘Azīz from Abū ‘Ubayd said: I heard Abū ‘Ubayda say: The one who claims that there is any other tongue than Arabic in the Qur’ān has opposed God with his words. He stands in need of His exalted utterance: “We have made it an Arabic Qur’ān.”

Abū ‘Ubayd said: it has been related from Ibn ‘Abbās, Mujāhid, ‘Akrima and others beside them, in many words (*ḥurūf*): that “sijjīl”, “al-miškāt”, “al-yamm”, “aṭ-ṭūr”, “abārīq”, “istabraq” and the like, is from other than Arabic.

These were better acquainted with explication (*ta’wīl*) than Abū ‘Ubayda. Even so, they followed one mode of thought, and this one (Abū ‘Ubayda) followed a different one.

Both are correct, God willing.

Which means: the origin of these words is not [found] in the Arabic language. [There are] those [who] said with regard to the origin: then the Arabs articulated it in their own language and thereby Arabicized it, and it became Arabic by virtue of its having become Arabicized. So it is Arabic at present but of foreign provenance.

This verdict is accepted as perfectly true by both factions. (Al-Jawālīqī 52-53)

An example of the latter case of a supposedly imported word having been naturalized by virtue of its similarity to Arabic word patterns is the entry *al-nūra* ‘one of the stones [which is] burnt [to produce] lime’ (al-Jawālīqī 389, note 8). The word is attributed to the name of the one—likely a non-Arab—believed to have introduced it to the Arabs.

*Al-nūra* was said not to be Arabic in its origin. [Yet] Its derivation resembles that of the Arabic derivations. Then it was claimed that it was called so since the first to have used it was a woman said to be Lahā Nūra. The Arabs used it in [their] poetry of old... (p. 389-90)

#### 4.2.2.1.4 Reflections over the place and function of loan words in the language

He shows an apologetic tone in having recourse to an argument attributing the Arabs’ usage of foreign vocabulary to “confusion” on their part regarding the true meaning of certain terms; the point he makes is, however, important, since it relates to one of the common motives underlying such usage, viz. the adding of phonological effect through rhyme to poetry, as well as of semantic originality—a function that could otherwise be filled by the use of *ġarīb* (cf. section 2.2.5). He also remarks on the irregular construction of Persian nouns (which word class is the largest contributor of loan words to Arabic) as compared with its counterpart in Arabic:

Abū ‘Umar al-Jarmī said: The Arabs may have confused [certain] foreign terms when they brought them into their [own] language. On the authority of Abū l-Mahdī, he said:

They say *šanbiḍ* to me, and not *mušanbiḍ*...

There is no *zūḍ* that makes my Lord hasten  
And the *bustān* in my breast is lofty, great...

By *šaḍbiḍ* we intend *šūn buḍī*. [By] *zūḍ*, ‘a‘jal and [by] *bustān huḍ*.



He said: When it was related to you regarding foreign [vocabulary] (*al-'a'jamīya*), the divergence [that] was characteristic of it, you shall certainly not see it as delirium (*taḥlīṭan*). The Arabs have caused confusion in the matter, and speak about it in an inconsistent manner since it is not [an integral] part of their [own] speech. So when they [alternately] censured (*i'tanafū*) it and spoke it, they muddled it.

Al-Farrā' has said: The Persian noun was constructed in whichever [random] manner, without any restricting rules, when it had not originated from an Arabic construction.

Abū Ḥātim recounted: Ra'ūba ibn al-'Ajjāj and the pure-spoken ones (*al-fuṣaḥā*), such as al-A'shā and others - : they seized on one of the foreign words for rhyme so as to make [their verse] sound original, but they did not make use of [what would be] considered unusual, nor inflect it, derive any verbs from it, or discard the unusual original. They might scoff at it as coming from the speech of an enemy:

I am the *bāk* Arab

That is to say: The [one] untainted by imperfection. 'Ajjāj said: As you saw in [the case of] *al-mulā' al-bardajā* : it was thought [to mean] 'captivity' since they had heard in Persian *bardah*. He thus intended a rhyme by it. (p. 54-58)

An element of humorous irony is evident in the latter example of Abū Ḥātim, since *bāk* is an Arabization of Persian *pāk* 'pure.'

#### 4.2.2.1.5 Ordering of the entries

The organization of the book's contents is alphabetical, but only by the words' first letters:

We have ordered this book according to the alphabetical (*mu'jam*) letters so as to ease (*yashula*) its craving (*marām*) and perfect its composition. (Al-Jawālīqī 60)

There is a chapter for words beginning with each of the letters for which the supposed foreign vocabulary items exist, totaling 26 chapters. The two missing chapter

s are for *dād* and *ṭā'* since, “these letters were not articulated (*yanṭuqu*) as the Arabs [articulated them]” (al-Jawālīqī 268).

This simple arrangement is not the most practical, as it obliges one to scan the whole chapter in order to ascertain whether or not a given word is represented. Yet it should be noted that, as with *Kitāb al-‘Ayn*, this dictionary was meant for scholars, not laymen, and the traditional practice was for them to memorize such works (Haywood 39-40). Also of significance is the limited number of entries in each chapter, averaging only 28, which makes it manageable to consult, with the exception, perhaps, of the larger ones, *bā'*, *alif*, *qāf*, *sīn*, *mīm*, whose numbers range from 67 to 50 entries (refer to Appendix 1 for some further statistics).

#### 4.2.2.1.6 Transcription of supposed loan words

Regarding the method used in *al-Mu‘arrab* for transcribing supposed loan vocabulary from Persian containing sounds not found in the Arabic alphabet, there is indication that al-Jawālīqī simply chooses the letter in Arabic that corresponds closest to the foreign sound. An example is Persian لگام ‘bridle’, which is attributed to the language and transcribed in *al-Mu‘arrab* as لغام *laḡām*, and listed under its Arabicized form اللجام *lajām* (al-Jawālīqī 348); Likewise Persian گردن *gardan* ‘neck’ is transcribed as کردن *kardan*, and listed under الكرد *al-kard* (al-Jawālīqī 327).

#### 4.2.2.2 Al-Jawālīqī’s treatment of *samā'*

Al-Jawālīqī uses the canonical corpora of the Arabic linguistic tradition as the basis for his compilation, and draws on its recognized linguistic authorities. He cites transmitted material including opinions of the learned regarding the purported foreignness of words, and both he and his sources is diligent in supplying the corresponding chain of transmission.

The question of foreign vocabulary in Arabic can be subsumed under the general activity of distinguishing and defining accepted usage in the language, a discipline that in the Arabic linguistic tradition was based firmly on *samā'* (cf. section 2.5.1.2.1). Both the author of and the sources cited in *al-Mu‘arrab* also naturally base their discussions about individual supposed loan words on this same criterion of attested usage, which principle is scrupulously

adhered to throughout the work as attested by the often numerous citations given for each entry.

### 4.2.2.3 Analytical concepts and methodological approaches

The conceptual tools that al-Jawālīqī himself and his cited sources make use of are generally based on al-Khalīl's and Sībawayhi's theories on the morphological and phonological characteristics of Arabic words. Also included are rationalizing arguments about the existence of certain forms. The technical vocabulary used to describe linguistic phenomena is consistent throughout, which may be surprising considering the span of several centuries that the references cover. This fact may be seen as yet another testimony to the enduring legacy of Sībawayhi's *Kitāb*, whose conceptual and terminological precedents remained largely unchallenged throughout the period in question.

### 4.2.2.4 Themes

The natural diversity and cultural richness of the conquered territories induced the borrowing of terminology in a range of fields which would have been largely unfamiliar to the Arabs (Asbaghi EALL Vol. III: 581), of which the following are selected examples from *al-Mu'arrab*. One may note that the entries' contents are often encyclopedic in nature, going into great detail of the themes covered; Fleisch has accordingly identified it as an "explanatory lexicon" (Fleisch 490).

#### 4.2.2.4.1 Plants, fruits, animals

As one might expect, there are many names of plant and animal species that have been adopted as loan words. Examples include the following:

Abū Bakr ibn Durayd said: *al-baṭṭa* [duck]: this bird; not pure Arabic. *Al-butṭ* is with the Arabs the small ones, and the big ones are *'iwazza*... (p. 112)

Abū Bakr said: *samandar* [salamander]: they claimed it is an animal. He said: I do not reckon it to be genuine Arabic. (p. 244)

Al-Azharī said: As to *al-šibitt* [dill] for this well-known herb, it is Arabicized. He said: I heard that the people of Bahrayn say *sibitt* for it...Its original in Persian is *šiwid* and it contains another dialect (*luḡa*) *sibiṭṭ* with *ṭā'*. (p. 257)

*Al-fustuq* [pistachio]: the singular is *fustuqa*. Arabicized Persian. It is a well-known fruit. They have given word to it... (p. 286)

#### 4.2.2.4.2 Christians

There are several references to Christian subjects, which is not surprising considering that they made up the largest religious minority in the empire.

##### *The Gospel*

*'Injīl*: Arabicized foreign. Some of them said: It was Arabic, then they inflected it from *al-najl*, which is the appearance of water on the face of the earth and its extension (*'ittisā'*). And *najaltu al-šay'* when you extracted and revealed it. So [from] *injīl* is extracted knowledge and judgment. And it was said: It is *'if'īl* from *al-najl* which is its original (*'aṣl*). So *al-'injīl* is the root (*'aṣl*) of knowledge and wisdom. (p. 71)

##### *Two words for archbishop*

*Asquf*: The Christians: Arabicized foreign. And they said *asquff*... (p. 83)

*Miṭrān* is Christian. It is not in pure Arabic. (p. 363)

##### *Christian festivals*

*Al-bāḡūt* [corresponds to *bā'ūt* 'Easter']. Arabicized foreign... (p. 105)

*Al-dinḥ*: ...It is not pure Arabic, being Arabicized... (p. 192)

*Al-sullāq* [Ascension of Christ]... Foreign which the Arabs knew. (p. 244)

Al-Azharī said: *al-nasṭūrīya* [Nestorian]: A community of Christians... In Romaeon it is *nasṭūris*. (p. 378)

#### 4.2.2.4.3 Certain “un-Islamic” subjects

Al-Jawālīqī has no scruples about discussing subjects inimical to the principles of Islam. On the other hand, this is in keeping with the general scientific curiosity and openness that characterized Islam in its early period. There is also the well known tradition of Muḥammad bidding his followers to seek knowledge even unto China. The following are entries for words denoting a wine-label and drunkenness, respectively:

The names *al-’isfanṭu*, *al-’isfinṭu*, *al-’isfand* and *al-’isfind* are among wine labels. It has been told me of Ibn as-Sikkīt that he said: It is a Romaeon word that is Arabicized, and it is not about wine, but the juice of a grape. He said: The people of al-Shām call *al-’isfanṭ* for *al-rasaṭūn*. They cook and add spices to it, then leave it to age.

Ibn Qutayba has recited to us *al-’isfinṭ* and *al-’isfind*: wine. And Ibn Abī Sa’īd said: *al-’isfanṭ* and *al-’isfand*. They said: It is the most sublime wine and its purest specimen. Al-A’shā said:

As if the aged wine from al-Isfinṭ  
is mixed with fresh pure water... (Al-Jawālīqī 66)

Al-Aṣma‘ī: drunkenness *ṭabarzad*, *ṭabarzal* and *ṭabarzan*: Three Arabicized dialects. Its original in Persian is *tabarzad*, as if meaning: hewn from its sphere with an ax. And *al-tabar*: Ax in Persian. And therefrom is named *al-ṭabarzad* from dates, since its date palm is as if smitten with an ax. (p. 276)

#### 4.2.2.4.4 Prophets and names figuring in the Qur’ān

The rationale for accounting the names of the prophets recorded in the Qur’ān as foreign is the obvious fact of their having been unknown to the Arabs during the *jāhiliyya*. In reference to *mūsā* ‘Moses,’ Abū l-‘Alā’ neatly summarizes this historicity and the mechanism by which the prophets’ names found their way into the language, as quoted in *al-Mu‘arrab*:

## Identifying foreign vocabulary in Arabic

There was not known anyone called *mūsā* in the *jāhilīya*. Rather, this occurred in Islam when the Qur'ān was sent down, and the Muslims called their sons by the names of the prophets...for a blessing, so when they called someone *mūsā* there presented itself a foreign name...and it is with them as 'īsā is. (p. 350)

Following are the particulars al-Jawālīqī has included regarding the names appearing in the Qur'ān, generally deemed to be Arabicized foreign:

The names of the prophets are all foreign, such as 'ibrāhīm, 'ismā'īl, 'ishāq, 'ilyās, 'idrīs, 'isrā'īl, and ayyūb, except for four names which are: 'ādam, ṣāliḥ, šu'ayb, and muḥammad. (p. 61)

'Irmīyā'u is the name of the prophet... (p. 69)

'Iblīs: Not in Arabic...Some of them say: it is Arabic... (p. 71)

'Āzar: Name of the father of Ibrāhīm. Abū Ishāq said: There is not disagreement among the people that Ibrāhīm's father's name was Tārah, [whereas] what is [stated] in the Qur'ān shows that his name is 'āzar. And it was said that 'āzar is a derogation (*ḍamm*) in their language (*luḡa*), as though it were: at fault (*maḥṭī'*)... (p. 76)

Ibn al-A'rābī said: It was mentioned from Ka'b...: The names of the Prophet in the former books, *muḥammad*, *'aḥmad* and *ḥimyāṭā* mean: sacred protector. (p. 170)

*Al-laysa'* and *lūṭ* are names of prophets... (p. 347)

*Mūsā* is the name of the prophet... Its original in Hebrew is *mūšā*. Thus *mū-* is water and *-šā* is tree, since he was found by the water and tree. (p. 350)

*Mīkā'īl*, Ibn 'Abbās said: *jabrā'īl* and *mīkā'īl*: *jabr*: Servant, as in your speech: 'Abd Allāh and 'Abd al-Raḥmān. It meant that 'īl was the name of God, the exalted, and the name of the angel *jabr* and *mīkā*, so they were related to God, the exalted. The exegetes have not differed on this. (But) the Qur'ān reciters differed in their recitals:

So some of them read *mīkā'il*. And some *mīkāl*. And others *mīkā'il*. And Ibn Muḥayṣin recited *mīka'il*. Like *mīka'il*. Al-Ḥarbī said: Abū 'Umar informed me from al-Kisā'ī saying: *jibrīl* and *mīkā'il* are names that the Arabs could not have known, [so when] they came they Arabicized them. (p. 375)

*Nūḥ* is the name of a prophet... (p. 378)

*Ya'qūb*: Name of the prophet... Also *yūsuf*, *yūnus*, *yūša'* and *al-yasa'*: All of them are foreign. (p. 403)

As to the entry on *mūsā* above, it can be noted that the etymological rationale given, “*mū-* is water and *-šā* is tree,” would not be accepted as valid according to modern etymological theory since it does not explain the etymology of the two supposed constituents.

#### 4.2.2.4.5 Application of miscellaneous grammatical rules

Following are some examples of rules of the language being applied or referred to in the entries of *al-Mu'arrab*:

Ibn al-Anbārī said: in *jahannam* are two utterances. Yūnus ibn Ḥabīb said, or most of the grammarians: *jahannam* is the name of the fire that God torments with in the afterlife. It is foreign... It is said, it is Arabic, and does not conform either to being rendered feminine or definite. It is related from Ru'ba that he said: *rakīya jihannām*: far-reaching pit. ...

...the absence of conjugation suggests that it is Arabicized foreign. (p. 155)

*Al-simsār*... Its verb is *al-samsara*: It was Arabicized. In the *ḥadīṭ* from Qays ibn Abī Gharaza: “We were called *al-samāsira*, so the Prophet called us by the most beautiful [of names], he said: “In the assemblage of the merchants”. ...

Abū Naṣr said, *simsār* is a man: [the one] who obeys him. (p. 249)

–*Andal*, goodness, does not have a root/original in the language. But they say: without *ṣandal*: when it was solid. (p. 268)

The letters *ḍād* and *ẓā'* do not have chapters. This is because these letters were not articulated (*yanṭuqu*) as the Arabs. (p. 268)

In the *ḥadīṭ* of al-Sha'abī occurs: He said to someone: You are bringing us these solid (*qasīya*) *ḥadīṭs* and taking them from us *ṭāzaja*: and *al-ṭāzaja*: unmixed purity. It is a conjugation of *tāzah*. (p. 277)

*Al-far'ana* is derived from *fir'awn* ('Pharaoh'). It is not in Arabic. (p. 294)

*Al-qabā'*, some said: It is Arabicized Persian. It is said: It is Arabic. Its derivation is from *al-qabw* and it is: gathering and collection. (p. 310)

#### 4.2.2.4.6 Personages, peoples, languages, common names

Following is a varied collection of entries from *al-Mu'arrab* referring to people, languages etc.

'*Asbaḍ*: Abū 'Ubayda said: Name of one of the Persian (*kisrā*) chiefs (*qā'id*) of Bahrayn, Persian...

Another than Abū 'Ubayda said: '*abīd asbaḍ* is a people of the family of Bahrayn, worshiping *al-barāḍīn* ('work horses').

'*Asbaḍ* is Persian, which Ṭarafa Arabicized. Its original is *asb* and it recalls *al-barāḍīn*.... (p. 86-7)

I read by way of ('*alā*) Abū Zakarīyā': It is said: '*iskandar* and '*askannadar*... He said: Abū 'Alā' mentioned it, so he told me: It is a foreign word, it has no cognate in Arabic speech. (p. 89)

Abū Ḥātim said: Al-Aṣma'ī said: It is said: *buḥtu naṣṣaru*, the one who devastated Jerusalem. (p. 128)

*Al-jawq*: a group of people. (p. 142)



*Fāris*: the name of this generation of people. Arabicized Persian... (p. 291)

*Julandā'*: Name of a king of 'Umān. Al-A'shā used it... (p. 155)

*Al-ḥayqār* (or *al-ḥīqār*): One of the Persian kings... Khālīd related *ḥīqār*, which is a man, and it is said: a tribe. (p. 169)

*Al-ḥusrawānī*: Al-Ḥarīr al-Raqīq al-Ḥasan al-Ṣan'a. It is attributed to the great ones of the descendants of the Persian kings (*al-'akāsira*). The Arabs have given word to it... (p. 183)

*Al-diryāq*: a language in *al-tiryāq*. It is Arabicized Romaeon. (p. 190)

*Dāwud* ('David): foreign. (p. 197)

*Dāhir*: name of the king of *al-daybul*. Foreign, which Jarīr used in his poetry... (p. 198)

*Rūmānis* in Romaeon. (p. 206)

*Al-rūm*: This is a race of people. Foreign. The Arabs of old have given word to it. And the Qur'ān uses it. (p. 211)

*Ṭā'ūs*: foreign. The Arabs of old have given word to it, and they called [people] by it. (p. 273)

*Hāmān*: A foreign name. *-fa'lān* is not from *hawwamtu* and not from *hāma yahīmu*. And do you not see that if you were to make the *alif* appended and the *nūn* a part of the [original] root in *hāmān*, like *sābāt*, you would not decline that either. (p. 398)

*Yahūd*: Arabicized foreign. They were related to *Yahūdā* ibn *Ya'qūb*. They called it *al-yahūd*, which was Arabicized with *dāl*.

## Identifying foreign vocabulary in Arabic

It was said that it is Arabic, and it was called *yahūdī* for repentance at some time. ... (p. 405)

## Conclusion

---

This study has presented developments in the Arabic linguistic tradition that relate to the question of identifying and distinguishing foreign vocabulary in the language. The connection between the linguistic tradition and the issue of loan words in Arabic from other languages has been made explicit wherever possible. Otherwise the connection has been made implicit, albeit within reason.

The present thesis has linked landmark developments in Arabic linguistics, particularly in grammar, on the one hand, to the issue of discerning the criteria for identifying loan words, on the other, thus demonstrating the close relationship that exists between these two conceptual bodies and the dependence of the latter's progress on the developments of the former—with al-Khalīl ibn Aḥmad and Sībawayhi figuring as key players in this regard. Al-Jawālīqī's dictionary *al-Muʿarrab* is a practical and comprehensive summary of statements and views concerning words considered to have been imported into Arabic, and the recognized sourcebook to this day on the subject; a fact which has evoked the need for special attention to be accorded it here. The work sheds light on issues related to the application of grammatical rules, and even the creation of the rules themselves, inasmuch as the morphological characteristics of a significant percentage of the words listed in *al-Muʿarrab* deviate noticeably from the norms outlined in grammar, and thus provide an ideal testing ground for determining the validity and effectiveness of these norms. The juxtaposition of conceptual theory and pragmatic methodology in Arabic grammar with actual examples of supposed foreign vocabulary items as listed in the entries of *al-Muʿarrab* which has been made throughout this study has served to clarify the nature of the link between the two entities. The breadth of theoretical considerations that have been found pertinent to this exercise—as witnessed in the multiplicity and variety of themes treated in the sizeable yet limited bibliography—testifies to its basic role in the wider scheme of linguistics.

The numerous translations that have been made from the chapters and entries in al-Jawālīqī's work, which has heretofore only been available in the original Arabic, can serve as a springboard for its more widespread use by English speaking scholars, and as a possible starting point for further in-depth studies on foreign vocabulary and related topics in Arabic as

## Conclusion

much as in other languages. The field of comparative linguistics can likewise benefit potentially from the themes covered herein.

The issue of loan words can be viewed as a portal to investigating the complex and largely obscure discipline of etymology. In this regard, the continued systematic mapping of the tangled network of stronger and weaker connections that make up the present knowledge base in etymology—closely related to the subject of foreign vocabulary as it is—must be counted on to open up significant new paths for exploring still-unexamined features in the linguistic system in general. The otherwise conservative field of Arabic lexicography which has been superficially approached here can likewise benefit from a more systematic treatment of the parameters upon which it is based viz. attested usage (*samāʿ*), which has been dealt with in our study. Sībawayhi's integral system—variously touched upon in this thesis—for the evaluation and categorization of instances of the latter, can even today serve as a model for such studies in etymology.

## Appendix 1

### Selected statistics for the entries in *al-Mu‘arrab*

#### Languages referred to

A rough count of the number of word entries (out of the ca. 719 in total) in which a word is attributed to or somehow connected with the different languages that are mentioned in *al-Mu‘arrab* is shown in Table 2. It should be noted that words are often attributed to more than one language by the different sources that are cited, or even by the same source.

*Table 1: Number of entries with attributions to the different languages.*

Arabic transcription	English equivalent	Approximate no. of entries
<i>fārsīya</i>	Persian/Pahlavi	257
<i>rūmīya</i>	Romaeon (a “dialect” of Greek)	32
<i>nabaṭīya</i>	Nabatean	21
<i>suryānī</i>	Syriac	19
<i>‘ibrānīya</i>	Hebrew	11
<i>ḥabašīya</i>	Abyssinian/Ethiopic	8
<i>hindīya</i>	Indian	1
<i>qibṭīya</i>	Coptic	1

It often occurs that a word is simply labeled or referred to as *‘a‘jamī* ‘foreign’, without any indication of which language it may have originated from. Approximately 131 of the entries use this designation. The term *‘a‘jamī* was sometimes used by Muslim linguists to actually designate Persian, which was considered the foreign language *par excellence* in relation to Arabic.

For a discussion about the implications of the designations given by the Arabic philologists to the different languages with reference to Qur'ānic words, refer to Jeffery *Foreign Words*, pp. 11-41.

On the next page are given statistics for the number of entries per chapter in *al-Mu'arrab*.

## Number of entries per chapter in *al-Mu‘arrab*

alif	64
bā’	67
tā’	18
ṭā’	1
jīm	44
ḥā’	14
ḥā’	24
dāl	43
ḍāl	1
rā’	20
zāy	30
sīn	51
šīn	20
ṣād	20
(ḍād	0)
(ṭāy	0)
ẓāy	24
‘ayn	11
ḡayn	3
fā’	35
qāf	64
kāf	40
lām	7
mīm	50
nūn	29
wāw	4
hā’	23
yā’	12

## Chapter ranking according to number of entries

1. bā’	67
2. alif	64
2. qāf	64
4. sīn	51
5. mīm	50
6. jīm	44
7. dāl	43
8. kāf	40
9. fā’	35
10. zāy	30
11. nūn	29
12. ḥā’	24
12. ẓāy	24
14. hā’	23
15. rā’	20
15. šīn	20
15. ṣād	20
18. tā’	18
19. ḥā’	14
20. yā’	12
21. ‘ayn	11

Total number of entries: 719 (approx.)

Average number of entries per chapter (out  
of 26 chapters): 27.7

## Appendix 2

### Miscellaneous turns of speech used in reference to *samā*‘ ‘attested data’

In light of the pivotal role in all aspects of the linguistic enterprise of establishing clear lines of authority through *samā*‘, it will inspire confidence in the user of *al-Mu‘arrab* to observe the close attention given to this aspect of the dictionary. Table 1 shows examples of the diversified turns of speech used in citing references by both al-Jawālīqī and his sources throughout the work.

Table 2: Miscellaneous turns of speech used in citing references in *al-Mu‘arrab*

Arabic text in transcription	Translation	page: line
<i>qara’tu ‘an ... ‘an ... qāla:</i>	I read from [so and so] from [so and so] [who] said:	61:5
<i>qara’tu ‘alā ...</i>	I read by way of [so and so]...	89:4
<i>wa-qad takallamat bi-hi l-‘arab ‘alā wujūh, fa-qālū...</i>	the Arabs pronounced it in [different] ways, saying...	61:6
<i>ruwā ‘anna ... qāla</i>	it has been related that [so and so] said...	61:8
<i>ruwā li-...</i>	it has been related of [so and so]...	61:10
<i>ruwā lī ‘an...</i>	it has been told me from...	66:4
<i>ruwā ‘an ... ‘annahā qālat</i>	it has been related from [so and so] that she said...	85:1
<i>yuqāl</i>	‘it is said...	62:3
<i>anšada ...</i>	[so and so] recited...	64:2
<i>qāla ba‘ḍuhum</i>	some said...	75:8
<i>qāla ba‘ḍu ahl al-‘ilm</i>	certain learned ones said...	65:7
<i>law qāla qā’il</i>	even if someone should say...	66:1



## Appendix 2

<i>lakāna qawlan</i>	as has been remarked	66:2
<i>qāla qawm min ahl al-luġa</i>	a group of linguists said...	67:11
<i>qāla ba‘ḍu ahl al-luġa...</i>	some of the linguists said...	91:7
<i>takallamat bi-hi l-‘arab</i>	the Arabs gave word to it	68:7
<i>qad takallamat bi-hi l-‘arab</i>	the Arabs have given word to it	74:4-5
<i>qad takallamat bi-hi l-‘arab qadīman</i>	the Arabs of old have given word to it	73:7
<i>qad jā’a fī l-ši‘r al-faṣīḥ</i>	it has appeared in the eloquent poem...	69:7
<i>ḥukiya ‘an...</i>	[such and such] was reported from [so and so]...	70:5
<i>fa-ammā... fa-...</i>	As to ..., it is....	73:1
<i>‘aḥaḍū ḍālika min...</i>	they took it from...	73:3
<i>za‘ama ... ‘anna</i>	[so and so] claimed that...	76:1
<i>laysa bayna l-nās ḥilāf ‘anna</i>	There is not disagreement among the people that...	76:11 -77:1
<i>mimmā jā’a ‘alā lafẓihi min alfāẓ al-‘arab</i>	[a word] that came from one of the Arabs’ dialects is...	81:1
<i>min ḍālika...</i>	Among them [is]...	81:9
<i>min ḍālika qawluhum li-...</i>	Among them is their saying for	79:7
<i>min ḍālika qawluhum fī-...</i>	Among them is their saying for...	81:3
<i>fī-hi luġāt</i>	There are [several] dialect [forms] of it...	82:1-2
<i>aḥbarnā ... ‘an... qāla...</i>	[so and so] informed us from [so and so] [who] said...	88:1
<i>hākaḍā ḍakarahu ... fa-qāla lī...</i>	[so and so] mentioned it [thus], so he said to me...	89:5
<i>sami‘tu al-‘arab taqūlu li-...</i>	I heard the Arabs say [such and such term] for [such and such thing]	90:1
<i>ḍakara ... ‘annahu...</i>	[so and so] mentioned that it is...	92:1
<i>qawluhum.../qawluhu...</i>	their utterance.../his utterance...	95:5/ 101:3
<i>hākaḍa fassarahu... ‘an...</i>	[so and so] explained it [thus] by way of [so and so]	117:6

# Bibliography

## *I. Primary sources*

Al-Jawālīqī. *al-Mu'arrab min al-kalām al-'a'jamī 'alā ḥurūf al-mu'jam*, Shākir, Aḥmad Muḥammad (ed.) 1969 C.E./1389 A.H.

Al-Jawālīqī. *Kitāb al-Mu'arrab min al-kalām al-'a'jamī 'alā ḥurūf al-mu'jam*, Sachau, E. (ed.) Leipzig, 1867

Al-Khalīl ibn Aḥmad, *Kitāb al-'Ayn*, li-Abī 'Abd al-Raḥmān al-Khalīl ibn Aḥmad al-Farāhīdī; taḥqīq Maḥdī al-Makhzūmī, Ibrāhīm al-Sāmarrā'ī. Baghdād: al-Jumhūrīyah al-'Irāqīyah, Wizārat al-Thaqāfah wa-al-Ilām, Dār al-Rashīd: Tawzī' al-Dār al-Waṭanīyah, 1980-1985

Sībawayhi, *Kitāb Sībawayhi*, Būlāq (1898-1900), reprint Baghdad (1965), Hārūn, A.S. (ed.) Cairo, 1968-77

## *II. Secondary sources*

Arberry, Arthur J. 1996. *The Koran Interpreted: A Translation*, Touchstone

Asbaghi, Asya. "Persian Loanwords," in: *Encyclopedia of Arabic Language and Linguistics*, 4 vols., Ed. by Versteegh, Kees, Leiden: Brill, Vol. III (580-84)

Baalbaki, Ramzi. 2008. *The Legacy of the Kitāb*, Leiden: Brill

—. 1998. "Kitāb al-'ayn and Jamharat al-lugha," in: Ryding, Karen C., *Early Medieval Arabic: Studies on al-Khalīl ibn Aḥmad*, Georgetown University Press (44-62)

Bergman, Elizabeth, M. "Orality," in: *Encyclopedia of Arabic Language and Linguistics*, 4 vols., Ed. by Versteegh, Kees, 2001-2006, Leiden: Brill (499-501)

## Bibliography

Bernards, Monique. "Ism," in: *Encyclopedia of Arabic Language and Linguistics*, 4 vols., Ed. by Versteegh, Kees, 2001-2006, Leiden: Brill (Vol. II: 424-29)

Bohas, G., Guillaume, J.-P., Kouloughli, D.E. 1990. *The Arabic Linguistic Tradition*, Routledge

Bussmann, Hadumod. 1996. *Routledge Dictionary of Language and Linguistics*, "Borrowing," Routledge (55-56)

Calder, Norman. "Tafsir from Tabari to Ibn Kathir: problems in the description of a genre, illustrated with reference to the story of Abraham," in: Hawting G.R., Shareef, Abdul-Kadar A. (Eds). 1993. *Approaches to the Qur'an*, Routledge (101-40)

Carter, Michael. 2006. "Foreign Vocabulary," in: *The Blackwell Companion to the Qur'an*, Rippin, Andrew (ed.) (120-39)

\_\_\_\_\_. 2004. *Sībawayhi*, I.B.Tauris

\_\_\_\_\_. "The Struggle for Authority: A Re-examination of the Baṣran and Kūfan debate," in: Edzard, Lutz and Nekroumi, Mohammed, eds. 1999. *Tradition and Innovation: Norm and Deviation in Arabic and Semitic Linguistics*, Wiesbaden: Harassowitz (55-70)

\_\_\_\_\_. "Another Khalīl: Courtier, Teacher and Sage," in: Ryding, Karen C. 1998. *Early Medieval Arabic: Studies on al-Khalīl ibn Aḥmad*, Georgetown University Press (16-43)

\_\_\_\_\_. "Sībawayhi," in: *Encyclopaedia of Islam*, Second Edition, Leiden: Brill, Vol. IX (524-31)

\_\_\_\_\_. 1973. "An Arab Grammarian of the Eighth Century A. D.: A Contribution to the History of Linguistics," *JAOS* Vol. 93, No. 2, (146-57)

Cowan, J. M. (ed.) 1994. *Arabic - English dictionary. The Hans Wehr dictionary of modern written Arabic*, Fourth Edition, Ithaca: Spoken Language Services

Crystal, David. 1993. *An Encyclopedic Dictionary of Language and Languages*, Blackwell

Daniels, Peter T. "Writings Systems," in: *The Handbook of Linguistics*, Ed. by Aronoff, Mark and Rees-Miller, Janie. 2001. Blackwell (43-80)

## Bibliography

Doughty, Charles Montagu. 2000. *Travels in Arabia Deserta*, Volume 1, Adamant Media Corporation

Enderwitz, S. "al-Shu'ūbiyya," in: *Encyclopaedia of Islam*, Second Edition, Leiden: Brill, Vol. IX (513-16)

Fleisch, H. "al-Djawālīkī or Ibn al-Djawālīkī, Abū Manṣūr Mawhūb b. Aḥmad b. Muḥ. b. al-Khaḍir." *Encyclopaedia of Islam*, Second Edition. Leiden: Brill, Vol. II (490)

Fox, A. "Orality," in: *Encyclopedia of Language and Linguistics*, Second Edition, Ed. by Keith Brown. 2006. Elsevier

Gustave E. von Grunebaum. 1944. "The Concept of Plagiarism in Arabic Theory," in: *Journal of Near Eastern Studies*, Vol. 3, No. 4 (Oct., 1944) (234-53)

Hamzé, Hassan Salam Bazzi. 2006. "Fi'l," in: *Encyclopedia of Arabic Language and Linguistics*, 4 vols., Ed. by Versteegh, Kees, 2001-2006, Leiden: Brill (90-96)

Hava, J.G. 1899. *Arabic-English Dictionary for the Use of Students*, Beirut: Catholic Press

Haywood, John A. 1965, *Arabic Lexicography*, Leiden: Brill

Heath, J. "Borrowing," in: *Encyclopedia of Language and Linguistics*, Asher, R. E. and Simpson, J. M. Y., eds., 1994, Pergamon Press, Vol. 1 (383-94)

Jeffery, Arthur. 1938. *The Foreign Vocabulary of the Qur'ān*, Baroda: Oriental Institute

Kopf, L. 1961. "The Treatment of Foreign Words in Medieval Arabic Lexicography," in: *Studies in Islamic History and Civilization*, Heyd, Uriel (ed.), Vol. IX, Jerusalem (191-205)

—. 1956. "Religious Influences on Medieval Arabic Philology," *Studia Islamica*, No. 5 (1956) (33-59)

## Bibliography

Lane, Edward William. 1968. *An Arabic-English Lexicon*, in Eight Parts, Beirut: Librairie du Liban

Lewis, Bernard. 2002. *The Arabs in history*, Sixth Edition, Oxford University Press

Madigan, Daniel. "Book," in: *Encyclopaedia of the Qur'án*, Ed. by McAuliffe, J.D., Leiden: Brill, 2001-2006, vol. I (242-51)

Maroti, Miklos. "Qiyās," in: *Encyclopedia of Arabic Language and Linguistics*, 4 vols., Ed. by Versteegh, Kees, 2001-2006, Leiden: Brill (11-14)

*Merriam-Webster Online Dictionary*. "Nationalism." 2009. Merriam-Webster Online. 5 May 2009 <<http://www.merriam-webster.com/dictionary/nationalism>>

Nadvi, Syed H. H., "Some Reflections on Islamic Education," in: *Encyclopaedic Survey of Islamic Culture: Growth and Development*, Taher, Mohamed (ed.), 1997, Anmol Publications, V.1-V.10 reprinted in 2003 (pp. 138-42). Accessed online 27 March 2009. <<http://books.google.com/books?id=ijdXih05lAYC&hl=no>>

Nasr, Seyyed Hossein and Leaman, Oliver. 1996. *History of Islamic Philosophy*, Routledge

Partridge, Eric. 1963. *The Gentle Art of Lexicography as Pursued and Experienced by an Addict*, A. Deutsch

Rippin, Andrew. "Foreign Vocabulary," in: *Encyclopaedia of the Qur'án*, Ed. by McAuliffe, J.D., Leiden: Brill, 2001-2006, Vol. III (226-37)

Rodinson, M. "Adjām," in: *Encyclopaedia of Islam*, Second Edition, Leiden: Brill, Vol. I (205-6)

Rodwell, J. M. 1876. *The Koran*. Accessed online 30 April 2009 at <<http://www.quranbrowser.com/>>

## Bibliography

Ryding, Karen C. 1998. *Early Medieval Arabic: Studies on al-Khalīl ibn Aḥmad*, Georgetown University Press

Seidensticker, Tilman. 2006. "Lexicography: Classical Arabic," in: *Encyclopedia of Arabic Language and Linguistics*, 4 vols., Ed. by Versteegh, Kees, 2001-2006, Leiden: Brill (30-37)

Sellheim, R. "al-Khalīl ibn Aḥmad," in: *Encyclopaedia of Islam*, Second Edition, Leiden: Brill, Vol. IX (962-64)

Shahīd, Irfan. 1998. Foreword in: Ryding, Karen C. 1998. *Early Medieval Arabic: Studies on al-Khalīl ibn Aḥmad*, , Georgetown University Press (vii-xi)

Stetkevych, Joseph. 1970. *The Modern Arabic Literary Language: Lexical and Stylistic Developments*, Georgetown University Press.

Suleiman, Yasir. "Arabiyya," in: *Encyclopedia of Arabic Language and Linguistics*, 4 vols., Ed. by Versteegh, Kees, 2001-2006, Leiden: Brill, Vol. I (173-78)

Talmon, Rafi. *From the History of the Study of Qur'ānic syntax*, in: *Tradition and Innovation: Norm and Deviation in Arabic and Semitic Linguistics*, Edzard, Lutz and Nekroumi, Mohammed, eds. 1998, Wiesbaden: Harrassowitz (71-77)

Versteegh, C. H. M. 1993. *Arabic Grammar and Qur'anic Exegesis in Early Islam*, Leiden: Brill

Versteegh, Kees. 1996. "Linguistic Attitudes and the Origin of Speech," in: *Understanding Arabic: Essays in Contemporary Arabic Linguistics in Honor of El-Said Badawi*, Elgibali, Alaa (ed.), 1996, The American University in Cairo Press

Versteegh, Kees, Versteegh C.H.M. 2006. "Learning Arabic in the Islamic World," in: *Handbook for Arabic language teaching professionals in the 21st century*, Wahba, Kassem M., Taha, Zeinab A., England, Liz, Routledge (3-12)

—. 2001. *The Arabic language*. Edinburgh University Press

## Bibliography

—. 1997. *Landmarks in Linguistic Thought III: The Arabic Linguistic Tradition*, Routledge

Westermarck, Edward. 1924. *The Origin and Development of the Moral Ideas*, Second Edition, London: Macmillan

## Abstract

This study presents developments in the Arabic linguistic tradition that relate to the question of identifying and distinguishing foreign vocabulary in the language. The connection between the linguistic tradition and the issue of loan words in Arabic from other languages has been made explicit wherever possible. Otherwise the connection is made implicitly.

The thesis links landmark developments in Arabic linguistics, particularly in grammar, on the one hand, to the issue of discerning the criteria for identifying loan words, on the other, thereby demonstrating the close relationship that exists between these two conceptual bodies and the dependence of the latter's progress on the developments of the former; al-Khalīl ibn Aḥmad and Sībawayhi figure as key players in this regard. Al-Jawālīqī's dictionary *al-Mu'arrab*, of which I have made a case study, gives a practical and comprehensive summary of statements and views concerning words considered to have been imported into Arabic, and is the recognized sourcebook to this day on the subject.

The numerous translations that have been made from the chapters and entries in al-Jawālīqī's work—which have heretofore only been available in the original Arabic—will serve as a potential springboard for its more widespread use by English speaking scholars, and as a possible starting point for further in-depth studies on foreign vocabulary and related topics in Arabic as much as in other languages.

A central feature of the thesis is the issue of loan words, which can be seen as a portal to investigating the complex discipline of etymology. The latter subject ties in with another theme treated here, viz. Arabic lexicography. Also treated are issues related to grammar—notably *qiyas* 'analogy' as applied by Sībawayhi. The latter's integral system for the evaluation and categorization of attested usage (*samāʿ*) has, moreover, been examined in the context of identifying foreign vocabulary in the language.